STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 142090

TO: Phuong Bui **Location: REM/2A15/2C18** **Art Unit: 1638** **Thursday, January 13, 2005** **Case Serial Number: 10/629953**	**From: Toby Port** **Location: Biotech-Chem Library** **REM-1A59** **Phone: 571-272-2523** **toby.port@uspto.gov**

Search Notes

Examiner Bui,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 13:59:30 ; Search time 3026 Seconds
(without alignments)
9876.765 Million cell updates/sec

Title: US-10-629-953-3
Perfect score: 632
Sequence: 1 ctcctccaagactccacaag.....taaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	610.8	96.6	691	8	AK072440	AK072440 Oryza sat
2	349.2	55.3	403	8	OSU16255	U16255 Oryza sativ
3	324.4	51.3	145419	8	AP004876	AP004876 Oryza sat
4	182.2	28.8	289	6	AR246100	AR246100 Sequence
5	149.6	23.7	263	6	AR246253	AR246253 Sequence
6	145.6	23.0	286	6	AR251557	AR251557 Sequence
7	96	15.2	324	8	BT004745	BT004745 Arabidops
8	96	15.2	605	8	AF053302	AF053302 Arabidops
9	68.2	10.8	913	8	AB050390	AB050390 Brassica
c 10	65.4	10.3	22758	8	AP004508	AP004508 Lotus cor
11	64.4	10.2	923	8	AK104192	AK104192 Oryza sat
12	64.4	10.2	945	8	AK070545	AK070545 Oryza sat
13	64.4	10.2	1141	8	AK059936	AK059936 Oryza sat
c 14	62.4	9.9	82896	8	ATT2K12	AL590346 Arabidops
c 15	62.4	9.9	103534	8	ATT5E8	AL391712 Arabidops
16	61	9.7	498	8	BT005847	BT005847 Arabidops
17	61	9.7	730	8	AF053303	AF053303 Arabidops
18	61	9.7	757	8	AY085896	AY085896 Arabidops
c 19	55.4	8.8	7218	6	I66494	I66494 Sequence 14
20	55.2	8.7	2000	6	AX655393	AX655393 Sequence
c 21	54.4	8.6	313846	1	AE017234	AE017234 Mycobacte
22	53	8.4	1293	6	BD180561	BD180561 Highly th
23	52.4	8.3	333	8	BT004696	BT004696 Arabidops
24	52.4	8.3	538	8	AK117904	AK117904 Arabidops
25	52.4	8.3	1074	6	AR105098	AR105098 Sequence
26	52.4	8.3	1074	8	AF244702	AF244702 Zea mays
27	51.4	8.1	273285	1	AE017304	AE017304 Thermus t
28	51.2	8.1	70162	2	AC026568	AC026568 Homo sapi
29	51	8.1	4725	1	AF142435	AF142435 Frankia s
30	51	8.1	314100	1	SCO939106	AL939106 Streptomy
31	50.2	7.9	110000	8	CR382131_34	Continuation (35 o
32	49.2	7.8	795	8	AY224434	AY224434 Oryza sat
33	49.2	7.8	795	8	AY224524	AY224524 Oryza sat
34	49.2	7.8	1152	8	AK101599	AK101599 Oryza sat
35	49.2	7.8	131049	8	AP005391	AP005391 Oryza sat
36	49.2	7.8	160089	8	AP004643	AP004643 Oryza sat
37	48.8	7.7	3443	8	AK120294	AK120294 Oryza sat
38	48.8	7.7	135216	8	AC069145	AC069145 Oryza sat
c 39	48.8	7.7	144724	8	AC078948	AC078948 Oryza sat
40	48.8	7.7	300029	8	AE017104	AE017104 Oryza sat
41	48.4	7.7	348257	1	BX640425	BX640425 Bordetell
c 42	48	7.6	100670	2	AP004682	AP004682 Oryza sat
c 43	48	7.6	148207	2	AP004697	AP004697 Oryza sat
44	48	7.6	170701	8	AP003510	AP003510 Oryza sat
c 45	48	7.6	301846	1	AE016913	AE016913 Chromobac

ALIGNMENTS

RESULT 1
AK072440
LOCUS AK072440 691 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023108D18, full insert sequence.
ACCESSION AK072440
VERSION AK072440.1 GI:32982463
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 691)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,

Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Direct Submission

JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .691
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023108D18"

ORIGIN

Query Match 96.6%; Score 610.8; DB 8; Length 691;
Best Local Similarity 99.7%; Pred. No. 8e-103;
Matches 612; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 101

Qy         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 161

Qy        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 221

Qy        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 281

Qy        241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        282 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGATA 341

Qy        301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 401

Qy        361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 461

Qy        421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        462 TGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 521

Qy        481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 581

Qy        541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 641

Qy        601 GGTTTCTGCGTGTA 614
              ||||||||||||||
Db        642 GGTTTCTGCGTGTA 655
```

RESULT 2
OSU16255
LOCUS OSU16255 403 bp mRNA linear PLN 10-JAN-1997
DEFINITION Oryza sativa clone ST1 sucrose-regulated mRNA, 3'-end sequence.
ACCESSION U16255
VERSION U16255.1 GI:576770
KEYWORDS .
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 403)
AUTHORS Tseng,T.C., Tsai,T.H., Lue,M.Y. and Lee,H.T.
TITLE Identification of sucrose-regulated genes in cultured rice cells using mRNA differential display
JOURNAL Gene 161 (2), 179-182 (1995)
MEDLINE 95394352
PUBMED 7665075
REFERENCE 2 (bases 1 to 403)
AUTHORS Lee,H.-t.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1994) Hung-tu Lee, National Tsing-Hua University, Life Science, No. 101, Sec. 2, Ken-Fu Rd., Hsinchu, Taiwan, R.O.C., 300
COMMENT On Nov 28, 1994 this sequence version replaced gi:571441.
FEATURES Location/Qualifiers
source 1. .403
/organism="Oryza sativa"
/mol_type="mRNA"
/strain="Taiwan 5"
/db_xref="taxon:4530"
/clone="SI1"
/cell_type="suspension cultured cell"
misc_feature 1. .403
/note="induced by sucrose"
/evidence=experimental
polyA_site 403

```
                     /note="11 A nucleotides"
ORIGIN

  Query Match             55.3%;  Score 349.2;  DB 8;  Length 403;
  Best Local Similarity   99.2%;  Pred. No. 1.5e-54;
  Matches  351;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        261 CCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGC 320
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CCGCGAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGC 61

Qy        321 TATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTGTGATGCAAGTA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTGTGATGCAAGTA 121

Qy        381 GACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATATTGCACTTTTGGTATGGTAAT 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATATTGCACTTTTGGTATGGTAAT 181

Qy        441 CTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGTTTGGTCATGCGCGGAAAATG 500
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGTTTGGTCATGCGCGGAAAATG 241

Qy        501 TGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGCCTAATGTGCTCTCCTCATTT 560
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TGGCCTCGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGCCTAATGTGCTCTCCTCATTT 301

Qy        561 TGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTTGGTTTCTGCGTGTA 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTTGGTTTCTGCGTGTA 355


RESULT 3
AP004876
LOCUS       AP004876               145419 bp    DNA     linear   PLN 28-JUL-2004
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
            PAC clone:P0470G10.
ACCESSION   AP004876
VERSION     AP004876.3  GI:46390400
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
            clone:P0470G10
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 145419)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     On Apr 14, 2004 this sequence version replaced gi:34850224.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'.  A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene
            predicted by a single gene prediction program is also classified as
            a probable 'hypothetical' protein and is included as a
            miscellaneous feature of the sequence.
            The orientation of the sequence is from SP6 to T7 of the PAC clone.
            This sequence of P0470G10 clone has an overlap with OJ1004_H01
            (DDBJ: AP004038) clone at 5' end and with OJ1008_F08 (DDBJ:
            AP004017) clone at 3' end. Detailed information on overlap and
            assembly quality together with annotation of this entry is
            available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1. .145419
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="2"
                     /clone="P0470G10"
     gene            complement(340. .3372)
                     /gene="P0470G10.1"
     misc_feature    complement(340. .3372)
                     /gene="P0470G10.1"
                     /note="probably inactive due to 5' exon missing in CDS
                     pseudogene, receptor protein kinase"
                     /pseudo
     gene            complement(4490. .5523)
                     /gene="P0470G10.2"
     misc_feature    complement(4490. .5523)
                     /gene="P0470G10.2"
                     /note="probably inactive due to 3' exon missing in CDS
                     pseudogene, Epstein-Barr virus EBNA-1-like protein"
                     /pseudo
     gene            complement(7419. .8108)
                     /gene="P0470G10.3"
     misc_feature    complement(7419. .8108)
                     /gene="P0470G10.3"
                     /note="probably inactive due to 3' exon missing in CDS
                     pseudogene, Hcr2-0A"
                     /pseudo
     gene            complement(8546. .10212)
                     /gene="P0470G10.4"
     mRNA            complement(join(<8546. .8551,9081. .9188,10147. .>10212))
                     /gene="P0470G10.4"
                     /note="start and end point are not identified"
     CDS             complement(join(8546. .8551,9081. .9188,10147. .10212))
                     /gene="P0470G10.4"
                     /note="predicted by GENSCAN etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD33295.1"
                     /db_xref="GI:50725764"
                     /translation="MARQVARAVVALVLRIEHKHNLRGGAWRRLILAIQDDVELLLLS
                     LLLVLVGGIPGPHLQ"
     gene            complement(11482. .14658)
                     /gene="P0470G10.5"
     mRNA            complement(join(<11482. .11634,11714. .11941,12441. .12545,
                     12620. .12697,12780. .12854,13509. .13595,13688. .13755,
                     13868. .13955,14055. .14155,14265. .14385,14434. .14479,
                     14585. .>14658))
                     /gene="P0470G10.5"
                     /note="start and end point are not identified"
     CDS             complement(join(11482. .11634,11714. .11941,12441. .12545,
                     12620. .12697,12780. .12854,13509. .13595,13688. .13755,
```

```
                13868. .13955,14055. .14155,14265. .14385,14434. .14479,
                14585. .14658))
                /gene="P0470G10.5"
                /note="contains full-length cDNA(s): AK100016,AK058244"
                /codon_start=1
                /product="putative thylakoid-bound ascorbate peroxidase"
                /protein_id="BAD33296.1"
                /db_xref="GI:50725765"
                /translation="MRGTCRRLRRRGAGAPARGRGRWSGYGTPAPSQLSHPYEQVRLG
                WHDSGTYDKNIEEWPQRGGADGSLRFDAELSHGANAGLINALKLIQPIKDKYPGITYA
                DLFQLASATAIEEAGGPKIPMKYGRVDVTAAEQCPPEGRLPDAGPRVPADHLREVFYR
                MGLDDKEIVALSGAHTLGRSRPDRSGWGKPETKYTKDGPGEPGGQSWTVEWLKFDNSY
                FKDIKEQRDQDLLVLPTDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAHAKLSDLGAK
                FDPPEGFSLDDEPAVEEKDPEPAPAPAAAPPPPPVEEKKEAEPTPVPVTVGAAVASSP
                ADDNNGAAPQPEPFVAAKYSYGKKELSDSMKQKIRAEYEGFGGSPDKPLQSNYFLNIM
                LLIGGLAFLTSLLGS"
gene            16256. .16777
                /gene="P0470G10.6"
mRNA            <16256. .>16777
                /gene="P0470G10.6"
                /note="start and end point are not identified"
CDS             16256. .16777
                /gene="P0470G10.6"
                /note="predicted by GeneMark.hmm etc."
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="BAD33297.1"
                /db_xref="GI:50725766"
                /translation="MPKTMSPRREATEAPPPPTSVKVSKRTETRLSPGRAARGVETAQ
                QRPQEGSDARNANPTKAGLGFHTLSVACISTADTPLLHHHLNLRWRVGTTTPAPPPAS
                LPMPPPAVPHRVALACTSPASTSASLASCCAGHISMRHAGLRHWHPSLPAPPSVAATS
                QTRLASLHTGCAS"
gene            complement(18792. .19088)
                /gene="P0470G10.7"
misc_feature    complement(18792. .19088)
                /gene="P0470G10.7"
                /note="hypothetical ORF
                predicted by GlimmerM
                this category is not included in IRGSP standard"
gene            complement(20341. .21044)
                /gene="P0470G10.8"
mRNA            complement(<20341. .>21044)
                /gene="P0470G10.8"
                /note="supported by full-length cDNA(s): AK110958"
misc_feature    complement(20341. .21044)
                /gene="P0470G10.8"
                /note="contains full-length cDNA(s): AK110958
                non-coding transcript
                probably inactive due to including stop codon(s) in CDS"
gene            22104. .25416
                /gene="P0470G10.9"
mRNA            join(<22104. .22246,23442. .23556,24206. .24344,24782. .24813,
                24943. .25029,25192. .>25416)
                /gene="P0470G10.9"
                /note="start and end point are not identified"
CDS             join(22104. .22246,23442. .23556,24206. .24344,24782. .24813,
                24943. .25029,25192. .25416)
                /gene="P0470G10.9"
                /note="contains EST(s): C28326(C60709)"
                /codon_start=1
                /product="putative tetratricopeptide repeat
                (TPR)-containing protein"
                /protein_id="BAD33298.1"
                /db_xref="GI:50725767"
                /translation="MRGFLMHQLLDSSSSDDDDELILAAALIAQHQYDIDNAPRRRGG
                SVRSDNDETEKEETRVVTTDCPGTKPSDNAESLKHQGNRLAEEGKYHEALGKWEAALT
                LMPNNAILHEQKAQILLELGDAWRALTAATRATELDPLWPEGWVTLGRAQLNFGEPDS
                AILSFDKALAIEPDNNEAKSDRETAARLVKKRGQLHSSGLSANKRRFTVGEDSEKEKD
                DSQMEDVEKCKEDEEKERDDSQLENVEK"
gene            27139. .29341
                /gene="P0470G10.10"
mRNA            join(27139. .27341,27451. .27556,28514. .28536,28983. .29341)
                /gene="P0470G10.10"
                /note="supported by full-length cDNA(s): AK072440"
CDS             join(27225. .27341,27451. .27556,28514. .28536,28983. .29042)
                /gene="P0470G10.10"
                /note="contains EST(s): AU082888(C53889),AU068340(C20066)
                contains full-length cDNA(s): AK072440"
                /codon_start=1
                /product="transcriptional coactivator p15 (PC4) family
                protein-like"
                /protein_id="BAD33299.1"
                /db_xref="GI:50725768"
                /translation="MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNR
                RVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA"
gene            complement(30379. .31415)
                /gene="P0470G10.11"
mRNA            complement(join(<30379. .30401,31091. .>31415))

  Query Match           51.3%;  Score 324.4;  DB 8;  Length 145419;
  Best Local Similarity  99.7%;  Pred. No. 5.4e-50;
  Matches  325;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        289 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28980 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 29039

Qy        349 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29040 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 29099

Qy        409 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAG 468
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      29100 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAG 29159

Qy        469 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29160 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 29219

Qy        529 GCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGT 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29220 GCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGT 29279

Qy        589 TAATGGTTGGTTGGTTTCTGCGTGTA 614
              |||||||||||||||||||||||||
Db      29280 TAATGGTTGGTTGGTTTCTGCGTGTA 29305


RESULT 4
AR246100
LOCUS       AR246100                289 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 1459 from patent US 6476212.
ACCESSION   AR246100
VERSION     AR246100.1  GI:27293974
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 289)
  AUTHORS   Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.
  TITLE     Polynucleotides and polypeptides derived from corn ear
  JOURNAL   Patent: US 6476212-A 1459 05-NOV-2002;
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           28.8%;  Score 182.2;  DB 6;  Length 289;
  Best Local Similarity  79.6%;  Pred. No. 9.4e-24;
  Matches  230;  Conservative    0;  Mismatches   53;  Indels    6;  Gaps    1;

Qy         41 GAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCGG 94
              | |  |||||| |||| || || ||||| ||||||| ||||||      ||||| ||||
Db          1 GGAAGATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGG 60
```

```
Qy          95 CCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCG 154
               |||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||||
Db          61 CCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120

Qy         155 TCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCG 214
               | ||| ||||||||||||||| || |||| | ||  ||| |||||| |||| |||| |||
Db         121 TAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCG 180

Qy         215 TCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC 274
               ||||| | ||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         181 TCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATAT 240

Qy         275 AGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTAT 323
                 |||| | | ||||||||||||||| ||| |||||| ||||||||  ||
Db         241 CACTCCAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCCAT 289


RESULT 5
AR246253
LOCUS       AR246253                 263 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 1612 from patent US 6476212.
ACCESSION   AR246253
VERSION     AR246253.1  GI:27294127
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 263)
  AUTHORS   Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.
  TITLE     Polynucleotides and polypeptides derived from corn ear
  JOURNAL   Patent: US 6476212-A 1612 05-NOV-2002;
FEATURES             Location/Qualifiers
     source          1. .263
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           23.7%;  Score 149.6;  DB 6;  Length 263;
  Best Local Similarity   78.0%;  Pred. No. 9.7e-18;
  Matches  206;  Conservative    0;  Mismatches   51;  Indels    7;  Gaps    2;

Qy          41 GAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCGG 94
               | |  |||||| |||| || || | ||| |||||||| ||||||      ||||| ||||
Db           1 GGAAGATGTGGGGGAAAGGAAAGANGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGG 60

Qy          95 CCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCG 154
               |||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||||
Db          61 CCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120

Qy         155 TCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCG 214
               | ||| ||||||||||||||| || |||| | ||  ||| |||||| |||| |||| |||
Db         121 TAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCG 180

Qy         215 TCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC 274
                |||| | |||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         181 NCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA- 239

Qy         275 AGCTCCCAATGGATCAGTGGAAGA 298
                  ||| | | ||||||||||||
Db         240 TCATCCAATTAGATCAGTGGAAGA 263


RESULT 6
AR251557
LOCUS       AR251557                 286 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 6916 from patent US 6476212.
ACCESSION   AR251557
VERSION     AR251557.1  GI:27299431
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.
  TITLE     Polynucleotides and polypeptides derived from corn ear
  JOURNAL   Patent: US 6476212-A 6916 05-NOV-2002;
FEATURES             Location/Qualifiers
     source          1. .286
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           23.0%;  Score 145.6;  DB 6;  Length 286;
  Best Local Similarity   78.4%;  Pred. No. 5.3e-17;
  Matches  189;  Conservative    0;  Mismatches   45;  Indels    7;  Gaps    1;

Qy          40 AGAAGAATGTGGCGGAAGGGGAACAAGCGGTTC-------GGCGGCGGCGGCGAGCCGGC 92
               || |  |||| | |||| || || ||||| |||       |||||| |||| ||||| ||
Db          44 AGGAAGATGTNGGGGAAAGGAAAGAAGCGTTTCGGGCGGTGGCGGCAGCGGTGAGCCCGC 103

Qy          93 GGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGT 152
               ||||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||
Db         104 GGCCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGT 163

Qy         153 CGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGT 212
               ||| ||| |||||||||||||||| || |||| | ||  ||| |||||| |||| |||| |
Db         164 CGTAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTT 223

Qy         213 CGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTAT 272
               ||||||| | ||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         224 CGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTAT 283

Qy         273 A 273
               |
Db         284 A 284


RESULT 7
BT004745
LOCUS       BT004745                 324 bp    mRNA    linear   PLN 22-FEB-2003
DEFINITION  Arabidopsis thaliana At5g09250 gene, complete cds.
ACCESSION   BT004745
VERSION     BT004745.1  GI:28466804
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 324)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 324)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-FEB-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
```

```
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J.,
            Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M.,
            Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yamada,K., Yu,G., Yuan,S., Davis,R.W.,  Theologis,A., and
            Ecker,J.R.

            Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .324
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="U23341"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS             1. .324
                     /note="putative transcriptional co-activator (KIWI)"
                     /codon_start=1
                     /product="At5g09250"
                     /protein_id="AAO44011.1"
                     /db_xref="GI:28466805"
                     /translation="MSSRGKRKDEDVRASDDESETHAPAKKVAKPADDSDQSDDIVVC
                     NISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKKGISLSVDQWNTLRNHAEDIEKA
                     LSDLS"
ORIGIN

  Query Match           15.2%;  Score 96;  DB 8;  Length 324;
  Best Local Similarity   66.3%;  Pred. No. 7.5e-08;
  Matches  138;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy        142 GACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAAC 201
              ||||||||||||||||  | | ||||| ||||| ||||| ||  | ||  |||  ||||||
Db        115 GACGATATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAAC 174

Qy        202 GGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGC 261
              || ||| |     | ||||| || |||||||| |  |||||||| |||||  | || |||
Db        175 GGGAAGATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGC 234

Qy        262 CGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCT 321
              ||||||||    ||     |||||||||||||| |  ||  || || |      | | |
Db        235 AAGAAAGGTATCTCTCTAAGCGTGGATCAGTGGAACACTCTTCGGAACCACGCAGAGGAT 294

Qy        322 ATAGATGAGGCCATCAAGGAGAATGCGT 349
              || ||  ||||| ||   ||   | | |
Db        295 ATCGAGAAGGCCCTCTCTGACCTTTCTT 322


RESULT 8
AF053302
LOCUS       AF053302                605 bp    mRNA    linear   PLN 03-JUL-2001
DEFINITION  Arabidopsis thaliana putative transcriptional co-activator (KIWI)
            mRNA, complete cds.
ACCESSION   AF053302
VERSION     AF053302.1  GI:2997683
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Cormack,R.S., Hahlbrock,K. and Somssich,I.E.
  TITLE     Isolation of putative plant transcriptional coactivators using a
            modified two-hybrid system incorporating a GFP reporter gene
  JOURNAL   Plant J. 14 (6), 685-692 (1998)
  MEDLINE   98346011
   PUBMED   9681033
REFERENCE   2  (bases 1 to 605)
  AUTHORS   Cormack,R.S., Hahlbrock,K. and Somssich,I.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-1998) Biochemistry, Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Cologne 50829, Germany
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /map="between KG31 and nga249"
     gene            1. .605
                     /gene="KIWI"
     CDS             97. .420
                     /gene="KIWI"
                     /codon_start=1
                     /product="putative transcriptional co-activator"
                     /protein_id="AAC08574.1"
                     /db_xref="GI:2997684"
                     /translation="MSSRGKRKDEDVRASDDESETHAPAKKVAKPADDSDQSDDIVVC
                     NISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKKGISLSVDQWNTLRNHAEDIEKA
                     LSDLS"
ORIGIN

  Query Match           15.2%;  Score 96;  DB 8;  Length 605;
  Best Local Similarity   66.3%;  Pred. No. 7.5e-08;
  Matches  138;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy        142 GACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAAC 201
              ||||||||||||||||  | | ||||| ||||| ||||| ||  | ||  |||  ||||||
Db        211 GACGATATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAAC 270

Qy        202 GGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGC 261
              || ||| |     | ||||| || |||||||| |  |||||||| |||||  | || |||
Db        271 GGGAAGATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGC 330

Qy        262 CGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCT 321
              ||||||||    ||     |||||||||||||| |  ||  || || |      | | |
Db        331 AAGAAAGGTATCTCTCTAAGCGTGGATCAGTGGAACACTCTTCGGAACCACGCAGAGGAT 390

Qy        322 ATAGATGAGGCCATCAAGGAGAATGCGT 349
              || ||  ||||| ||   ||   | | |
Db        391 ATCGAGAAGGCCCTCTCTGACCTTTCTT 418


RESULT 9
AB050390
LOCUS       AB050390                913 bp    mRNA    linear   PLN 14-SEP-2001
DEFINITION  Brassica rapa BcKELP mRNA for putative transcriptional coactivator,
            complete cds.
ACCESSION   AB050390
VERSION     AB050390.1  GI:13699909
KEYWORDS    .
SOURCE      Brassica rapa
  ORGANISM  Brassica rapa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Matsushita,Y., Deguchi,M., Youda,M., Nishiguchi,M. and Nyunoya,H.
  TITLE     The tomato mosaic tobamovirus movement protein interacts with a
            putative transcriptional coactivator KELP
  JOURNAL   Mol. Cells 12 (1), 57-66 (2001)
  MEDLINE   21445625
   PUBMED   11561731
REFERENCE   2  (bases 1 to 913)
  AUTHORS   Matsushita,Y. and Nyunoya,H.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Yasuhiko Matsushita, Tokyo University of
            Agriculture and Technology, Gene Research Center; 3-5-8 Saiwai-cho,
            Fuchu, Tokyo 183-8509, Japan (E-mail:ymatsu@cc.tuat.ac.jp,
            Tel:81-42-367-5855, Fax:81-42-367-5855)
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Brassica rapa"
                     /mol_type="mRNA"
                     /strain="S9 homozygote"
                     /db_xref="taxon:3711"
                     /clone="MIP102 (=BcKELP)"
                     /clone_lib="lambda GEX5 Brassica rapa flower bud cDNA"
                     /note="vector:pGEX-Bc2"
     gene            1. .913
                     /gene="BcKELP"
     CDS             14. .511
                     /gene="BcKELP"
                     /codon_start=1
                     /product="putative transcriptional coactivator"
                     /protein_id="BAB41214.1"
                     /db_xref="GI:13699910"
                     /translation="MEEESKAKIEETVREILKESDMTEMTEFKVRNLASERLGIDLSD
                     KSHKAFVRGIVKSFLEEVESKQQQQQDKEEEEEEEERAKEGNKEFDDDGDLIICRLS
                     DKRRVTIQEFRGKSLVSIREYYKKDGKELPSSKGISLTDEQWSTFKKNIPAIEAAVKK
                     MESRV"
     polyA_signal    869. .874
                     /gene="BcKELP"
                     /evidence=not_experimental
ORIGIN

  Query Match             10.8%;  Score 68.2;  DB 8;  Length 913;
  Best Local Similarity   57.9%;  Pred. No. 0.01;
  Matches  121;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              ||  |||| |||| || || |   |  | | ||| | || ||||| ||| || | | |
Db        286 CGATGACGGCGATCTCATCATTTGCAGGCTGTCGGATAAGAGGAGAGTGACGATTCAGGA 345

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              |  |  || |||    | ||   |||| | ||||  ||| |||| ||||||||    ||
Db        346 GTTTAGAGGAAAGAGTTTGGTTTCCATCAGAGAGTATTACAAGAAAGACGGCAAAGAGCT 405

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
               ||       ||||| |||    |  ||   || || |||   |   | | | | |||||
Db        406 TCCTTCTTCTAAAGGAATAAGCTTAACAGACGAACAATGGTCAACGTTCAAGAAAAATAT 465

Qy        315 CAAAGCTATAGATGAGGCCATCAAGGAGA 343
              |||||| || |  ||  ||||| | |
Db        466 TCCAGCTATCGAAGCTGCTGTCAAGAAAA 494


RESULT 10
AP004508/c
LOCUS       AP004508               22758 bp    DNA     linear   PLN 22-JUL-2003
DEFINITION  Lotus corniculatus var. japonicus genomic DNA, chromosome 1,
            clone:LjT03K03, TM0036c, complete sequence.
ACCESSION   AP004508
VERSION     AP004508.1  GI:17736875
KEYWORDS    HTG.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Sato,S., Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T. and Tabata,S.
  TITLE     Structural Analysis of a Lotus japonicus Genome. I. Sequence
            Features and Mapping of Fifty-six TAC clones which cover the 5.4 Mb
            Regions of the Genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 22758)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
            Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1. .22758
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /chromosome="1"
                     /clone="LjT03K03"
                     /clone_lib="LjT library"
                     /note="TM0036c, a part of TAC clone:TM0036~synonym: Lotus
                     japonicus"
ORIGIN

  Query Match             10.3%;  Score 65.4;  DB 8;  Length 22758;
  Best Local Similarity   73.0%;  Pred. No. 0.033;
  Matches   84;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy        158 CCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCG 217
              | ||||| ||||||||||||||||||  | ||| ||| |||| | || | | |||| ||||
Db      22653 CGCAGATTTCGAAGAACAGGAGGGTTTCTGTGAGGAACTGGCAGGGGAGGATCGTGGTCG 22594

Qy        218 ACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTAT 272
              |||| ||||| || || |  || |||||||||     | || ||    ||||| |
Db      22593 ACATTCGCGAATTTTATGTCAAAGACGGCAAGCAAATGCCTGGGAAGAAAGGTTT 22539


RESULT 11
AK104192
LOCUS       AK104192               923 bp    mRNA    linear   PLN 24-JUL-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:006-303-D08, full
            insert sequence.
ACCESSION   AK104192
VERSION     AK104192.1  GI:32989401
KEYWORDS    FLI_CDNA; oligo capping.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
  MEDLINE   22752273
   PUBMED   12869764
REFERENCE   2  (bases 1 to 923)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
```

Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Direct Submission

JOURNAL Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .923
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="006-303-D08"

ORIGIN

Query Match 10.2%; Score 64.4; DB 8; Length 923;
Best Local Similarity 59.2%; Pred. No. 0.05;
Matches 129; Conservative 0; Mismatches 86; Indels 3; Gaps 1;

```
Qy        136 GCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACC 195
              | ||| |  ||  || || ||  ||   | |||   |  |||||||||| |  | | |
Db        403 GACGAGGGAGACCTCATCCTCTGCCGCCTTTCGGCCAGGAGGAGGGTGACTCTTCAGGAG 462

Qy        196 TGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTC 255
              |  || ||||||    |  |   ||||||| |||| ||||   ||||||||||||   ||
Db        463 TTCAAGGGCAAGACGCTTCTGTCCATCCGAGAGTACTACTTCAAGGACGGCAAGGAGCTG 522

Qy        256 CCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATC 315
              ||||   |||||| |||    |  || | || ||||||  |   |  | ||   | |
Db        523 CCCG---CCAAAGGGATAAGTTTGACAGTTGAGCAGTGGGAAGCATTTCGTGATTCTGTG 579

Qy        316 AAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATC 353
                ||| ||||| || ||||| ||  ||  ||    |||
Db        580 CCAGCAATAGAGGATGCCATTAAAAAGCTTGGAGAATC 617
```

RESULT 12
AK070545
LOCUS AK070545 945 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023055M24, full insert sequence.
ACCESSION AK070545
VERSION AK070545.1 GI:32980569
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764

REFERENCE 2 (bases 1 to 945)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo;S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .945
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023055M24"

ORIGIN

Query Match 10.2%; Score 64.4; DB 8; Length 945;
Best Local Similarity 59.2%; Pred. No. 0.05;
Matches 129; Conservative 0; Mismatches 86; Indels 3; Gaps 1;

```
Qy        136 GCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACC 195
              | ||| |  ||  || || ||  ||   | |||   |  |||||||| |  | | |
Db        350 GACGAGGGAGACCTCATCCTCTGCCGCCTTTCGGCCAGGAGGAGGGTGACTCTTCAGGAG 409

Qy        196 TGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTC 255
              |  || ||||||    |  |   |||||| |||| ||||   ||||||||||||   ||
Db        410 TTCAAGGGCAAGACGCTTCTGTCCATCCGAGAGTACTACTTCAAGGACGGCAAGGAGCTG 469

Qy        256 CCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATC 315
              ||||    |||||| |||    |  || | || |||||| |   | |  | ||   | |
Db        470 CCCG---CCAAAGGGATAAGTTTGACAGTTGAGCAGTGGGAAGCATTTCGTGATTCTGTG 526

Qy        316 AAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATC 353
                ||| ||||| || ||||| ||  ||  ||    |||
Db        527 CCAGCAATAGAGGATGCCATTAAAAAGCTTGGAGAATC 564
```

RESULT 13
AK059936
LOCUS AK059936 1141 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:006-210-D10, full insert sequence.
ACCESSION AK059936
VERSION AK059936.1 GI:32969954
KEYWORDS FLI_CDNA; oligo-capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 1141)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,

```
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
            Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
            Yasunishi,A. and Hayashizaki,Y.
FEATURES             Location/Qualifiers
     source          1. .1141
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="006-210-D10"
ORIGIN

  Query Match             10.2%;  Score 64.4;  DB 8;  Length 1141;
  Best Local Similarity   59.2%;  Pred. No. 0.05;
  Matches  129;  Conservative    0;  Mismatches   86;  Indels    3;  Gaps    1;

Qy        136 GCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACC 195
              | ||| |  ||  || || ||  ||   | |||   |  |||||||| |  | | |
Db        428 GACGAGGGAGACCTCATCCTCTGCCGCCTTTCGGCCAGGAGGAGGGTGACTCTTCAGGAG 487

Qy        196 TGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTC 255
              |  || ||||||    |  |   ||||||| |||| ||||   ||||||||||||   ||
Db        488 TTCAAGGGCAAGACGCTTCTGTCCATCCGAGAGTACTACTTCAAGGACGGCAAGGAGCTG 547

Qy        256 CCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATC 315
              ||||    |||||| |||    |  || | || |||||| |   | |  | ||   | |
Db        548 CCCG---CCAAAGGGATAAGTTTGACAGTTGAGCAGTGGGAAGCATTTCGTGATTCTGTG 604

Qy        316 AAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATC 353
                ||| ||||| || ||||| ||  ||  ||    |||
Db        605 CCAGCAATAGAGGATGCCATTAAAAAGCTTGGAGAATC 642


RESULT 14
ATT2K12/c
LOCUS       ATT2K12                 82896 bp    DNA     linear   PLN 03-APR-2001
DEFINITION  Arabidopsis thaliana DNA chromosome 5, BAC clone T2K12 (ESSA
            project).
ACCESSION   AL590346
VERSION     AL590346.1  GI:13548323
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 82896)
  AUTHORS   Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,
            Rudd,S., Lemcke,K. and Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 82896)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2001) MIPS, at GSF/IBI, Ingolstaedter Landstr. 1,
            85764 Neuherberg, FRG, E-mail: schoof@gsf.de,kmayer@gsf.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://mips.gsf.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .82896
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
     misc_feature    1. .14902
                     /note="overlap to PAC MAH20, EMBL:AB006697; please refer
                     to this clone for annotation"
     gene            15981. .19300
                     /gene="At5g08660"
     CDS             join(15981. .16289,16433. .16746,16833. .16925,17020. .17110,
                     17305. .17398,17478. .17575,17767. .17836,18028. .18131,
                     18231. .18347,18458. .18534,18636. .18680,18763. .19300)
                     /gene="At5g08660"
                     /note="similarity to predicted proteins, Arabidopsis
                     thaliana"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAC35871.1"
                     /db_xref="GI:13548324"
                     /db_xref="TrEMBL:Q9C5B1"
                     /translation="MGSFCSKSLGINFGSEYSGSSVADDGREPDFGHSQPNGQTSLIV
                     PGMRQLMVKDVKEQNQLKDVFSFRERERAEDNFYDGIPTYTMAPSQKIRSAKSTQTAVS
                     KVTEASKLLGKAGLGRAKDVLDTLGSSMTDLSSGGFTSGVATKGNELGILAFEVANTI
                     VKSSNLIESLSKRNIEHLKGTILYSEGVQNLVSNDFDELLRLVAADKRQELQVFSGEV
                     VRFGNRSKDFQWHNLQRYFDRISKELTPQRQLKEDAVLVVDQLMVLVQYTAELYQELQ
                     VLYRLEKDYEQKRREEENSANSSKGDGLAILKTELKAQRKVVKSLKKKSLWSRGFEEV
                     MEKLVDIVHFLLLEIHNIFGGADDQPSKKGAAEYDKRLGPAGLALHYANIIVQIDTLV
                     ARASSITSNARDSLYQSLPPGIKLALRSKIKSFNVDKELSVTQIKDEMERTLHWLVPV
                     AGNTTKAHHGFGWVGEWANTGTDFTSKPSGGDILRIETLYHASKEKTEIYILGQIIWL
                     QHLVTKAKSDARGGPRLSSIKSPLDTTNQQLISEPLSVPIVTDEEQKMLQEASKRKRT
                     PCVSKSQDFDSEYSRARKCDPLSKSSEYFRGVRRSKSAAVKRYSSGFPLLDFAIDKEK
                     VLDVIDRVDVPRDYKALLKEGSLSF"
     exon            15981. .16289
                     /gene="At5g08660"
                     /number=1
     intron          16290. .16432
                     /gene="At5g08660"
                     /number=1
     exon            16433. .16746
                     /gene="At5g08660"
                     /number=2
     intron          16747. .16832
                     /gene="At5g08660"
                     /number=2
     exon            16833. .16925
                     /gene="At5g08660"
                     /number=3
     intron          16926. .17019
                     /gene="At5g08660"
                     /number=3
     exon            17020. .17110
                     /gene="At5g08660"
                     /number=4
     intron          17111. .17304
                     /gene="At5g08660"
                     /number=4
     exon            17305. .17398
                     /gene="At5g08660"
                     /number=5
     intron          17399. .17477
                     /gene="At5g08660"
                     /number=5
     exon            17478. .17575
                     /gene="At5g08660"
                     /number=6
     intron          17576. .17766
                     /gene="At5g08660"
                     /number=6
     exon            17767. .17836
                     /gene="At5g08660"
                     /number=7
     intron          17837. .18027
                     /gene="At5g08660"
                     /number=7
     exon            18028. .18131
                     /gene="At5g08660"
                     /number=8
     intron          18132. .18230
                     /gene="At5g08660"
                     /number=8
     exon            18231. .18347
```

```
                /gene="At5g08660"
                /number=9
     intron     18348. .18457
                /gene="At5g08660"
                /number=9
     exon       18458. .18534
                /gene="At5g08660"
                /number=10
     intron     18535. .18635
                /gene="At5g08660"
                /number=10
     exon       18636. .18680
                /gene="At5g08660"
                /number=11
     intron     18681. .18762
                /gene="At5g08660"
                /number=11
     exon       18763. .19300
                /gene="At5g08660"
                /number=12
     gene       19850. .22604
                /gene="At5g08670"
     gene       complement(join(19850. .19963,20204. .20515,20610. .20834,
                20918. .21076,21177. .21280,21360. .21477,21569. .21657,
                21738. .21813,22131. .22604))
                /gene="At5g08670"
     CDS        complement(join(19850. .19963,20204. .20515,20610. .20834,
                20918. .21076,21177. .21280,21360. .21477,21569. .21657,
                21738. .21813,22131. .22604))
                /gene="At5g08670"
                /note="strong similarity to H+-transporting ATP synthase,
                mitochondrial, rubber tree, PIR:S20504
                Contains ATP/GTP-binding site motif A (P-loop)
                AA231-238;ATP synthase alpha and beta subunits signature
                AA423-432"
                /codon_start=1
                /product="H+-transporting ATP synthase beta chain
                (mitochondrial)-like protein"
                /protein_id="CAC35872.1"
                /db_xref="GI:13548325"
                /db_xref="GOA:P83483"
                /db_xref="Swiss-Prot:P83483"
                /translation="MASRRVLSSLLRSSSGRSAAKLGNRNPRLPSPSPARHAAPCSYL
                LGRVAEYATSSPASSAAPSSAPAKDEGKKTYDYGGKGAIGRVCQVIGAIVDVRFEDQE
                GLPPIMTSLEVQDHPTRLVLEVSHHLGQNVVRTIAMDGTEGLVRGRKVLNTGAPITVP
                VGRATLGRIMNVLGEPIDERGEIKTEHYLPIHRDAPALVDLATGQEILATGIKVVDLL
                APYQRGGKIGLFGGAGVGKTVLIMELINNVAKAHGGFSVFAGVGERTREGNDLYREMI
                ESGVIKLGEKQSESKCALVYGQMNEPPGARARVGLTGLTVAEYFRDAEGQDVLLFIDN
                IFRFTQANSEVSALLGRIPSAVGYQPTLASDLGALQERITTTKKGSITSVQAIYVPAD
                DLTDPAPATTFAHLDATTVLSRQISELGIYPAVDPLDSTSRMLSPHILGEEHYNTARG
                VQKVLQNYKNLQDIIAILGMDELSEDDKLTVARARKIQRFLSQPFHVAEIFTGAPGKY
                VDLKENINSFQGLLDGKYDDLSEQSFYMVGGIDEVVAKAEKIAKESAA"
     exon       complement(19850. .19963)
                /gene="At5g08670"
                /number=1
     intron     complement(19964. .20203)
                /gene="At5g08670"
                /number=1
     exon       complement(20204. .20515)
                /gene="At5g08670"
                /number=2
     intron     complement(20516. .20609)
                /gene="At5g08670"
                /number=2
     exon       complement(20610. .20834)
                /gene="At5g08670"
                /number=3
     intron     complement(20835. .20917)
                /gene="At5g08670"
                /number=3
     exon       complement(20918. .21076)
                /gene="At5g08670"
                /number=4
     intron     complement(21077. .21176)
                /gene="At5g08670"
                /number=4
     exon       complement(21177. .21280)
                /gene="At5g08670"
                /number=5
     intron     complement(21281. .21359)
                /gene="At5g08670"
                /number=5
     exon       complement(21360. .21477)
                /gene="At5g08670"
                /number=6
     intron     complement(21478. .21568)
                /gene="At5g08670"
                /number=6
     exon       complement(21569. .21657)
                /gene="At5g08670"

  Query Match             9.9%;  Score 62.4;  DB 8;  Length 82896;
  Best Local Similarity   72.3%;  Pred. No. 0.12;
  Matches   81;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy        160 CAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGAC 219
              ||||||||| ||||| ||||| ||  | ||  |||  |||||||| ||| |     | |||
Db      77539 CAGATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAGATTTGGATTGAC 77480

Qy        220 ATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTA 271
              || || |||||||| |  |||||||| |||||  | || |||   |||||||
Db      77479 ATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAAGGTA 77428


RESULT 15
ATT5E8/c
LOCUS       ATT5E8                103534 bp    DNA     linear   PLN 29-AUG-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 5, BAC clone T5E8 (ESSA
            project).
ACCESSION   AL391712
VERSION     AL391712.1  GI:9955507
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 103534)
  AUTHORS   Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,
            Rudd,S., Lemcke,K. and Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 103534)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .103534
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
     gene            3395. .6144
                     /gene="T5E8_10"
     gene            complement(join(3395. .3451,3525. .3633,4011. .4306,
                     4412. .4685,4910. .5226,5275. .5397,5442. .5567,5635. .6144))
                     /gene="T5E8_10"
     CDS             complement(join(3395. .3451,3525. .3633,4011. .4306,
```

```
                4412. .4685,4910. .5226,5275. .5397,5442. .5567,5635. .6144))
                /gene="T5E8_10"
                /note="similarity to GCF2 fusion protein - Homo sapiens,
                EMBL:AB026911"
                /codon_start=1
                /product="putative protein"
                /protein_id="CAC05447.1"
                /db_xref="GI:9955508"
                /db_xref="TrEMBL:Q9FY92"
                /translation="MGTKQPRNYRRRCNDEIDGEDATAAAKPSSSDLYPAKPSSSDLY
                PWKRRKLPAIPTKNDDDAKGRDHAKSSWLLDLSRGDEFYTQEQMQQRLKEHSQDERPM
                EIVFNGERVRSIGMDGLSYVFVCPMPQTADDKMAALDRICNAYDERVNSRMPQEQPQM
                YAGRPEELAKEPDENLTMSIAAAESCPSAPVYKYASLQEISDFKSVFRNFMQGICVAF
                VCVSGAFISIKRLLSIFFLSQKNGYLITAIEDQMKVDGYSLIVEGDSSTDDESDCETS
                AYEEARDSLLQRADKIFSDASVVYSELSRVKSIFKRGARHPSPAFRAAYTSLTVPSMY
                SPYLRLELLRWDPLHQDVDFSDMNWHGLLFHSRIVCGSTPVCTNPNFVSELVKYVAVP
                ILHHRIVRCWDILSTRETRNVVAATSLVARYVFPSSEALAELSLAIHARLVEAIIAIS
                VPTWDPQVSKDVPNAPQVAAYRFGTSVRLMRNICMWKDVMELPVLEKLALSDLLFGKV
                LPHVRSIASESNIHDAVTKTERIVASLSGVWTGPSVTRTHSHLLQPLVDCTLTLGRIL
                EKKVCLGTGRYGNHWSRPQEDVDLQLTDPFSVRIKVI"
exon            complement(3395. .3451)
                /gene="T5E8_10"
                /number=1
intron          complement(3452. .3524)
                /gene="T5E8_10"
                /number=1
exon            complement(3525. .3633)
                /gene="T5E8_10"
                /number=2
intron          complement(3634. .4010)
                /gene="T5E8_10"
                /number=2
exon            complement(4011. .4306)
                /gene="T5E8_10"
                /number=3
intron          complement(4307. .4411)
                /gene="T5E8_10"
                /number=3
exon            complement(4412. .4685)
                /gene="T5E8_10"
                /number=4
intron          complement(4686. .4909)
                /gene="T5E8_10"
                /number=4
exon            complement(4910. .5226)
                /gene="T5E8_10"
                /number=5
intron          complement(5227. .5274)
                /gene="T5E8_10"
                /number=5
exon            complement(5275. .5397)
                /gene="T5E8_10"
                /number=6
intron          complement(5398. .5441)
                /gene="T5E8_10"
                /number=6
exon            complement(5442. .5567)
                /gene="T5E8_10"
                /number=7
intron          complement(5568. .5634)
                /gene="T5E8_10"
                /number=7
exon            complement(5635. .6144)
                /gene="T5E8_10"
                /number=8
gene            7652. .9648
                /gene="T5E8_20"
CDS             join(7652. .7793,7900. .8133,8215. .8308,8403. .8617,
                8698. .8837,8992. .9648)
                /gene="T5E8_20"
                /note="Contains Prokaryotic membrane lipoprotein lipid
                attachment site AA454-464"
                /codon_start=1
                /product="amino acid transport protein AAP2"
                /protein_id="CAC05448.1"
                /db_xref="GI:9955509"
                /db_xref="GOA:Q38967"
                /db_xref="TrEMBL:Q38967"
                /translation="MGETAAANNHRHHHHHGHQVFDVASHDFVPPQPAFKCFDDDGRL
                KRTGTVWTASAHIITAVIGSGVLSLAWAIAQLGWIAGPAVMLLFSLVTLYSSTLLSDC
                YRTGDAVSGKRNYTYMDAVRSILGGFKFKICGLIQYLNLFGIAIGYTIAASISMMAIK
                RSNCFHKSGGKDPCHMSSNPYMIVFGVAEILLSQVPDFDQIWWISIVAAVMSFTYSAI
                GLALGIVQVAANGVFKGSLTGISIGTVTQTQKIWRTFQALGDIAFAYSYSVVLIEIQD
                TVRSPPAESKTMKKATKISIAVTTIFYMLCGSMGYAAFGDAAPGNLLTGFGFYNPFWL
                LDIANAAIVVHLVGAYQVFAQPIFAFIEKSVAERYPDNDFLSKEFEIRIPGFKSPYKV
                NVFRMVYRSGFVVTTTVISMLMPFFNDVVGILGALGFWPLTVYFPVEMYIKQRKVEKW
                STRWVCLQMLSVACLVISVVAGVGSIAGVMLDLKVYKPFKSTY"
exon            7652. .7793
                /gene="T5E8_20"
                /number=1
intron          7794. .7899
                /gene="T5E8_20"
                /number=1
exon            7900. .8133
                /gene="T5E8_20"
                /number=2
intron          8134. .8214
                /gene="T5E8_20"
                /number=2
exon            8215. .8308
                /gene="T5E8_20"
                /number=3
intron          8309. .8402
                /gene="T5E8_20"
                /number=3
exon            8403. .8617
                /gene="T5E8_20"
                /number=4
intron          8618. .8697
                /gene="T5E8_20"
                /number=4
exon            8698. .8837
                /gene="T5E8_20"
                /number=5
intron          8838. .8991
                /gene="T5E8_20"
                /number=5
exon            8992. .9648
                /gene="T5E8_20"
                /number=6
gene            11840. .14398
                /gene="T5E8_30"
CDS             join(11840. .11995,12106. .12214,12604. .12841,12924. .12989,
                13090. .13135,13245. .13405,13482. .13617,13691. .13855,
                13929. .14056,14129. .14228,14348. .14398)
                /gene="T5E8_30"
                /note="similarity to sirtuin type 4 (SIRT4) - Homo
                sapiens, EMBL:AF083109"
                /codon_start=1
                /product="SIR2-family protein"
                /protein_id="CAC05449.1"
                /db_xref="GI:9955510"
                /db_xref="GOA:Q9FY91"
                /db_xref="TrEMBL:Q9FY91"
                /translation="MKLHQDKNRRESHERIFDRELTRSSSRERDRSEGEIEAKERELR
                PQRENKIDTSDFCHLVFNNNKKFWKKIKRKTEEISFHPPLTTSQDLFPSRSMYRPLQS
                GGNLVMLFKGCRRFVRTTCRVSIPGGSLGNESKAPPRFLRDRKIVPDADPPNMEDIHK
                LYRLFEQSSRLTILTGAGVSTECGIPDYRSPNGAYSSGFKPITHQEFTRSSRARRRYW
                ARSYAGWRRFTAAQPGPAHTALASLEKAGRINFMITQNVDRLHHRAGSDPLELHGTVY
                TVMCLECGFSFPRDLFQDQLKAINPKASWAEAIESIDHGDPGSEKSFGMKQRPDGDIE
                IDEKFWEEGFHIPVCEKCKGVLKPDVIFFGDNIPKERATQAMEVAKQSDAFLVLGSSL
                MTMSAFRLCRAAHEAGAMTAIVNIGETRADDIVPLKINARVGEILHRVLDVGSLSVPA
                L"
exon            11840. .11995
                /gene="T5E8_30"
                /number=1
```

```
     intron          11996. .12105
                     /gene="T5E8_30"
                     /number=1
     exon            12106. .12214
                     /gene="T5E8_30"
                     /number=2
     intron          12215. .12603
                     /gene="T5E8_30"
                     /number=2
     exon            12604. .12841
                     /gene="T5E8_30"
                     /number=3
     intron          12842. .12923
                     /gene="T5E8_30"
```

Query Match 9.9%; Score 62.4; DB 8; Length 103534;
Best Local Similarity 72.3%; Pred. No. 0.12;
Matches 81; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy        160 CAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGAC 219
              |||||||| ||||| ||||| ||  | ||  |||  |||||||| ||| |     | |||
Db      16869 CAGATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAGATTTGGATTGAC 16810

Qy        220 ATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTA 271
              || || |||||||| |  |||||||| |||||  | || |||   |||||||
Db      16809 ATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAAGGTA 16758
```

Search completed: January 11, 2005, 21:05:45
Job time : 3034 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 12:05:19 ; Search time 441 Seconds
(without alignments)
7522.980 Million cell updates/sec

Title: US-10-629-953-3
Perfect score: 632
Sequence: 1 ctcctccaagactccacaag.....taaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	632	100.0	632	3	AAZ50412	Aaz50412 Rice tran
	2	220.4	34.9	460	3	AAZ50414	Aaz50414 Wheat tra
	3	210.6	33.3	649	3	AAZ50411	Aaz50411 Corn tran
	4	182.2	28.8	289	10	ABX82999	Abx82999 Corn ear-
	5	149.6	23.7	263	10	ABX83152	Abx83152 Corn ear-
	6	146.6	23.2	310	6	ABL70693	Abl70693 Corn tass
	7	145.6	23.0	286	10	ABX88456	Abx88456 Corn ear-
	8	109.4	17.3	672	3	AAZ50413	Aaz50413 Soybean t
	9	96	15.2	471	3	AAC38126	Aac38126 Arabidops
	10	78.6	12.4	516	3	AAZ50416	Aaz50416 Vernonia
	11	76.4	12.1	939	3	AAZ50418	Aaz50418 Corn tran
	12	76	12.0	1089	3	AAZ50417	Aaz50417 Corn tran
	13	74	11.7	1063	3	AAC32760	Aac32760 Arabidops
	14	68.2	10.8	913	8	ACF03526	Acf03526 Brassica
	15	66.2	10.5	740	3	AAZ50419	Aaz50419 Soybean t
	16	61.6	9.7	498	3	AAZ50415	Aaz50415 Marigold
c	17	61	9.7	496	10	ADE81749	Ade81749 Arabidops
	18	61	9.7	757	3	AAC50198	Aac50198 Arabidops
	19	61	9.7	757	3	AAC34846	Aac34846 Arabidops
	20	55.2	8.7	2000	8	ADA71938	Ada71938 Rice gene
	21	52.4	8.3	1074	3	AAA76284	Aaa76284 Maize glu
	22	49.2	7.8	458	10	ADC72476	Adc72476 DNA Seq I
	23	49.2	7.8	458	10	ADC72595	Adc72595 DNA Seq I
	24	49.2	7.8	458	10	ADD17205	Add17205 DNA (SeqI
	25	49.2	7.8	458	10	ADK58353	Adk58353 Plant DNA
	26	48	7.6	1509	6	ABL39958	Abl39958 Synthetic
	27	48	7.6	1509	12	ADM73763	Adm73763 HIV-1 pol
	28	47.6	7.5	597	12	ADJ44561	Adj44561 Plant cDN
c	29	47.4	7.5	349980	6	ABQ81847	Abq81847 Bifidobac
	30	46.8	7.4	1848	9	ACF35993	Acf35993 Erythrocy
	31	46.8	7.4	1848	9	ACF35992	Acf35992 Erythrocy
	32	46.4	7.3	1509	3	AAA51610	Aaa51610 HIV synth
	33	46.4	7.3	1509	3	AAA51626	Aaa51626 HIV codon
	34	46.4	7.3	1509	6	AAL44553	Aal44553 HIV-1 p55
	35	46.4	7.3	1509	6	AAL44549	Aal44549 HIV-1 p55
	36	46.4	7.3	1509	6	ABL39954	Abl39954 Synthetic
	37	46.4	7.3	1509	12	ADM73759	Adm73759 HIV-1 pol
	38	46.4	7.3	1545	12	ADN36395	Adn36395 Human gen
	39	46.4	7.3	8186	12	ADN36401	Adn36401 Human gen
	40	46	7.3	1795	12	ADM47600	Adm47600 Polynucle
	41	45.8	7.2	954	11	ABD03958	Abd03958 Pseudomon
	42	45.8	7.2	1416	4	AAS51578	Aas51578 Pseudomon
	43	45.8	7.2	1416	8	ACA19447	Aca19447 Prokaryot
	44	45.8	7.2	1458	12	ADJ39777	Adj39777 Plant cDN
	45	45.8	7.2	1953	11	ABD03877	Abd03877 Pseudomon

ALIGNMENTS

RESULT 1
AAZ50412
ID AAZ50412 standard; cDNA; 632 BP.
XX
AC AAZ50412;
XX
DT 18-MAY-2000 (first entry)
XX
DE Rice transcription coactivator PC4(P15) type 1 cDNA.
XX
KW Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW rice; signal mediator; activator; Gal4/VP16; transgenic plant;
KW general transcription factor; GTF; transcription initiation complex;
KW immunological screening; detection; marker; clone rr1.pk0003.a12; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 46. .351
FT /*tag= a
FT /product= "Rice PC4(P15) type 1 transcription
FT coactivator"
FT /note= "Derived from clone rr1.pk0003.a12"
XX
PN WO200005377-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016479.
XX
PR 22-JUL-1998; 98US-0093687P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;
XX
DR WPI; 2000-182701/16.
DR P-PSDB; AAY44879.
XX
PT Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT polynucleotides and polypeptides.
XX
PS Claim 9; Page 33; 43pp; English.

```
XX
CC   The present sequence is the cDNA encoding rice Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone rr1.pk0003.a12, obtained from rr1 cDNA library prepared using
CC   rice root two week old developing seedlings. PC4 molecules functions as a
CC   signal mediator between activators like Gal4/VP16 and general
CC   transcription factors (GTFs) in a transcription initiation complex. The
CC   expression levels of PC4 can be manipulated and the functional properties
CC   of specific transcriptional activators can be modulated. The PC4 DNA can
CC   be used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 632 BP; 165 A; 139 C; 187 G; 141 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 632;  DB 3;  Length 632;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-154;
  Matches  632;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60

Qy         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120

Qy        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180

Qy        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240

Qy        241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300

Qy        301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360

Qy        361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420

Qy        421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480

Qy        481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540

Qy        541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600

Qy        601 GGTTTCTGCGTGTAAAAAAAAAAAAAAAAAAAA 632
              ||||||||||||||||||||||||||||||||
Db        601 GGTTTCTGCGTGTAAAAAAAAAAAAAAAAAAAA 632


RESULT 2
AAZ50414
ID   AAZ50414 standard; cDNA; 460 BP.
XX
AC   AAZ50414;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Wheat transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   wheat; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone wdk2c.pk015.g20; ss.
XX
OS   Triticum aestivum.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .293
FT                   /*tag= a
FT                   /product= "Wheat PC4(P15) type 1 transcription
FT                   coactivator"
FT                   /note= "Derived from clone wdk2c.pk015.g20"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44881.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 35; 43pp; English.
XX
CC   The present sequence is the cDNA encoding wheat Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone wdk2c.pk015.g20, obtained from wdk2c cDNA library, prepared
CC   from wheat developing kernels, 7 days after anthesis. PC4 molecules
CC   functions as a signal mediator between activators like Gal4/VP16 and
CC   general transcription factors (GTFs) in a transcription initiation
CC   complex. The expression levels of PC4 can be manipulated and the
CC   functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 460 BP; 117 A; 110 C; 139 G; 87 T; 0 U; 7 Other;

  Query Match             34.9%;  Score 220.4;  DB 3;  Length 460;
  Best Local Similarity   79.3%;  Pred. No. 2.7e-47;
  Matches  260;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy         72 CGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGA 131
              ||||||| ||||| ||||| ||||||||||||  || |||||| | |||||| ||||||||
Db         14 CGGCGGCCGCGGCCAGCCGCCGGCCAAGCGCCAGGCGGCCGGGAAGGACGGTCCCTCCGA 73

Qy        132 GAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCG 191
              |    |||||||||  |||||||| || ||||||||||||||| |||||||||  ||| |
Db         74 GGAAACCGACGACGGCATCGTCGTTGCGCAGATATCGAAGAACAAGAGGGTGGCCGTGAG 133

Qy        192 GACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGAC 251
              || ||||||||| |||||| | |||||||| ||||||||||||| ||||||||||||
Db        134 GAACTGGAACGGGAAGGTCATGGTCGACATGCGCGAGTTCTACGAAAAGGACGGCAAGAG 193

Qy        252 CCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAA 311
```

```
                ||||||   |||||||||||||  |||| ||||||||||||||||  ||||||||||||||
Db        194 CCTCCCGACCCGCAAAGGTATATCGCTCTCAATGGATCAGTGGAAAATACTGAGGGACAA 253

Qy        312 TATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTGTG 371
                ||| ||||||||||  ||||||||||||||||||   | |||||  ||      |  |
Db        254 CATCGAAGCTATAGACGAGGCCATCAAGGAGAACACTTGATCAGAAAAGCGGTTAAGGAN 313

Qy        372 ATGCAAGTAGACTAAGCCTACGTCTGTC 399
                ||   | | | | |||  |    ||  ||
Db        314 ATNTTACTGGGCAAAGTGTTGCTCATTC 341


RESULT 3
AAZ50411
ID   AAZ50411 standard; cDNA; 649 BP.
XX
AC   AAZ50411;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Corn transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   corn; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone cca.pk0020.d2; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             33. .344
FT                   /*tag=  a
FT                   /product= "Corn PC4(P15) type 1 transcription
FT                   coactivator"
FT                   /note= "Derived from clone cca.pk0020.d2"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44878.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 32; 43pp; English.
XX
CC   The present sequence is the cDNA encoding corn Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from the corn callus type II tissue (undifferentiated) from the cDNA
CC   clone cca.pk0020.d2. PC4 molecules functions as a signal mediator between
CC   activators like Gal4/VP16 and general transcription factors (GTFs) in a
CC   transcription initiation complex. The expression levels of PC4 can be
CC   manipulated and the functional properties of specific transcriptional
CC   activators can be modulated. The PC4 DNA can be used to create transgenic
CC   plants with altered PC4 levels, that would affect the level of
CC   transcription of specific genes in the plant. It is also used for
CC   immunological screening of cDNA libraries and to raise specific
CC   antibodies for detection. The DNA is used as probes and primers, for
CC   genetic and physical mapping of genes and as markers for traits linked to
CC   those genes
XX
SQ   Sequence 649 BP; 167 A; 147 C; 182 G; 153 T; 0 U; 0 Other;

  Query Match             33.3%;  Score 210.6;  DB 3;  Length 649;
  Best Local Similarity   77.2%;  Pred. No. 1.1e-44;
  Matches  271;  Conservative    0;  Mismatches   74;  Indels    6;  Gaps    1;

Qy         15 CACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGG 74
              |||  ||   |      |||      |  ||  |  ||||||  ||||  ||  ||  |||||  |||||
Db          2 CACGAGGTCCAGTCGAGTCTCAGCAAGGAAGATGTGGGGGAAAGGAAAGAAGCGTTTCGG 61

Qy         75 CGGCGGCGGC------GAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTC 128
              |||  ||||||      |||||  ||||||||||||||  |||||  |||||||  ||  |||||
Db         62 CGGTGGCGGCAGCGGTGAGCCCGCGGCCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTC 121

Qy        129 CGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGT 188
              |||   |||  ||  ||||  ||  ||||||  |||  ||||||||||||||||  ||  ||||  |  ||
Db        122 CGAATCCGCTGAAGACGGTACCGTCGTAGCCGAGATATCGAAGAACAAGAAGGTGTCCGT 181

Qy        189 GCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAA 248
                 |||  ||||||  ||||  ||||  ||||||||  |  ||||||||||||    ||||||||||
Db        182 TAGGAGCTGGAAAGGCAGGGTCTTCGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAA 241

Qy        249 GACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGA 308
              |||  ||||||  ||||||||||||||    ||||  |  |  ||||||||||||||||||||  |||
Db        242 GACTCTCCCCACCCGCAAAGGTATATCACTCCAATTAGATCAGTGGAAGATACTGAAGGA 301

Qy        309 CAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCC 359
              |||  ||||||||  |||  |||||||  |||  ||||  |||  |  ||||  ||||||
Db        302 CAACATCAAAGCCATAAATGAGGCAATCGAGGAAAATACATGATGGGAGCC 352


RESULT 4
ABX82999
ID   ABX82999 standard; cDNA; 289 BP.
XX
AC   ABX82999;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #1459.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
XX
PS   Example; SEQ ID NO 1459; 390pp; English.
XX
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
```

```
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 289 BP; 79 A; 71 C; 92 G; 47 T; 0 U; 0 Other;

  Query Match             28.8%;  Score 182.2;  DB 10;  Length 289;
  Best Local Similarity   79.6%;  Pred. No. 2e-37;
  Matches  230;  Conservative    0;  Mismatches   53;  Indels    6;  Gaps    1;

Qy          41 GAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCGG 94
               | |  |||||| |||| || || ||||| |||||||| ||||||      ||||| ||||
Db           1 GGAAGATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGG 60

Qy          95 CCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCG 154
               |||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||||
Db          61 CCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120

Qy         155 TCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCG 214
               | ||| |||||||||||||| || |||| | ||  ||| |||||| |||| |||| |||
Db         121 TAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCG 180

Qy         215 TCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC 274
               ||||| | ||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         181 TCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATAT 240

Qy         275 AGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTAT 323
                 |||| | | |||||||||||||||| ||| |||||| ||||||| ||
Db         241 CACTCCAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCCAT 289


RESULT 5
ABX83152
ID   ABX83152 standard; cDNA; 263 BP.
XX
AC   ABX83152;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #1612.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
XX
PS   Example; SEQ ID NO 1612; 390pp; English.
XX
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 263 BP; 67 A; 65 C; 88 G; 41 T; 0 U; 2 Other;

  Query Match             23.7%;  Score 149.6;  DB 10;  Length 263;
  Best Local Similarity   78.0%;  Pred. No. 6e-29;
  Matches  206;  Conservative    0;  Mismatches   51;  Indels    7;  Gaps    2;

Qy          41 GAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCGG 94
               | |  |||||| |||| || || | ||| |||||||| ||||||      ||||| ||||
Db           1 GGAAGATGTGGGGGAAAGGAAAGANGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGG 60

Qy          95 CCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCG 154
               |||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||||
Db          61 CCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120

Qy         155 TCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCG 214
               | ||| |||||||||||||| || |||| | ||  ||| |||||| |||| |||| |||
Db         121 TAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCG 180

Qy         215 TCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC 274
               ||||| | ||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         181 NCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA- 239

Qy         275 AGCTCCCAATGGATCAGTGGAAGA 298
                 ||| | | ||||||||||||||
Db         240 TCATCCAATTAGATCAGTGGAAGA 263


RESULT 6
ABL70693
ID   ABL70693 standard; cDNA; 310 BP.
```

```
XX
AC   ABL70693;
XX
DT   14-MAY-2002  (first entry)
XX
DE   Corn tassel-derived polynucleotide (cdps) SEQ ID NO:67.
XX
KW   Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW   inheritance; characteristic; growth; development; disease resistance;
KW   environmental adaptability; quality; yield; molecular marker;
KW   multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS   Zea mays.
XX
PN   US2001051335-A1.
XX
PD   13-DEC-2001.
XX
PF   16-APR-1999;   99US-00294093.
XX
PR   21-APR-1998;   98US-0082567P.
XX
PA   (LALG/) LALGUDI R V.
PA   (ITOL/) ITO L Y.
PA   (SHER/) SHERMAN B K.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2002-163647/21.
XX
PT   Novel purified corn tassel-derived polynucleotide useful for determining
PT   altered gene expression, to recover regulatory elements and to follow
PT   inheritance of desirable characteristics through hybrid breeding
PT   programs.
XX
PS   Claim 1; SEQ ID NO 67; 201pp; English.
XX
CC   The present sequence describes a purified corn tassel-derived
CC   polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC   selected from those given in ABL70627 to ABL76833. The cdps sequences
CC   encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC   can be used for determining altered gene expression, to recover
CC   regulatory elements and to follow inheritance of desirable
CC   characteristics through hybrid breeding programs. (I) are also useful in
CC   the evaluation, and alteration of desired characteristics associated with
CC   growth and development, disease resistance, environmental adaptability,
CC   quality and yield, and as molecular markers for studying inheritance of
CC   multigene traits in a plant breeding program. (I) can be used to produce
CC   a tassel-specific profile of gene transcription, a transcript image, to
CC   clone regulatory elements for use in transformation vectors, to express a
CC   polypeptide, to identify, isolate or extend identical or related corn
CC   tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC   hybridisation or amplification technologies, as query sequences to
CC   determine homology of known sequences, as probe for use in Southern or
CC   Northern hybridisation, and to identify the presence of and/or to
CC   determine the degree of similarity between two (or more) nucleic acid
CC   sequences
XX
SQ   Sequence 310 BP; 84 A; 77 C; 96 G; 50 T; 0 U; 3 Other;

  Query Match            23.2%;  Score 146.6;  DB 6;  Length 310;
  Best Local Similarity  71.2%;  Pred. No. 3.8e-28;
  Matches  220;  Conservative    0;  Mismatches   86;  Indels    3;  Gaps    2;

Qy         17 CAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCG 76
              | |||| ||    || |||   ||       | || | |||  |       |    ||||
Db          3 CCAGTCGAGTCTCAGCAAGGAAGATGTNGGGGAGGGGAAAGAAGCGTTTCGGCGGTGGCG 62

Qy         77 GCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCC-GGGGACGACGGGCCCTCCGAGAGC 135
              || |||||||||| ||||||||||||  |||||  |||||||| || ||||||||   |
Db         63 GCAGCGGCGAGCCCGCGGCCAAGCGCCAGGCCGCGAGGGGACGATGGCCCCTCCGAATCC 122

Qy        136 GCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACC 195
              || || |||| || |||||| ||| |||||||||||||| || |||  | ||  ||| |
Db        123 GCTGAAGACGGTACCGTCGTAGCCGAGATATCGAAGAACAAGAAGGTATCCGTTAGGAGC 182

Qy        196 TGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTC 255
              ||||| |||| |||| |||||||| | ||||||||||||   ||||||||||||| |||
Db        183 TGGAAAGGCAGGGTCTTCGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTC 242

Qy        256 CCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATC 315
              |||  |||||||||     ||  | |||||       |||||||| ||| |||| | |||
Db        243 CCCACCCGCAAAGTATATCA--TTCCAATTAGATCATGGAAGATATTGAAGGACNACATC 300

Qy        316 AAAGCTATA 324
              ||||| |||
Db        301 AAAGCCATA 309


RESULT 7
ABX88456
ID   ABX88456 standard; cDNA; 286 BP.
XX
AC   ABX88456;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #6916.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
XX
PS   Example; SEQ ID NO 6916; 390pp; English.
XX
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
```

```
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 286 BP; 69 A; 77 C; 92 G; 47 T; 0 U; 1 Other;

  Query Match            23.0%;  Score 145.6;  DB 10;  Length 286;
  Best Local Similarity  78.4%;  Pred. No. 6.8e-28;
  Matches  189;  Conservative    0;  Mismatches   45;  Indels    7;  Gaps    1;

Qy         40 AGAAGAATGTGGCGGAAGGGGAACAAGCGGTTC-------GGCGGCGGCGGCGAGCCGGC 92
              || |  |||| | |||| || || ||||| |||       ||||||  |||| ||||| ||
Db         44 AGGAAGATGTNGGGGAAAGGAAAGAAGCGTTTCGGGCGGTGGCGGCAGCGGTGAGCCCGC 103

Qy         93 GGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGT 152
              ||||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||
Db        104 GGCCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGT 163

Qy        153 CGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGT 212
              ||| ||| ||||||||||||||| || |||| | ||  ||| |||||| |||| |||| |
Db        164 CGTAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTT 223

Qy        213 CGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTAT 272
              ||||||| | ||||||||||||   |||||||||||||| ||||||  ||||||||||
Db        224 CGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTAT 283

Qy        273 A 273
              |
Db        284 A 284


RESULT 8
AAZ50413
ID   AAZ50413 standard; cDNA; 672 BP.
XX
AC   AAZ50413;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Soybean transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   soybean; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone sfl1.pk0008.a4; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             37. .354
FT                   /*tag= a
FT                   /product= "Soybean PC4(P15) type I transcription
FT                   coactivator"
FT                   /note= "Derived from clone sfl1.pk0008.a4"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44880.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 33-34; 43pp; English.
XX
CC   The present sequence is the cDNA encoding soybean Positive Cofactor 4
CC   (PC4) transcription coactivator, designated as PC4(P15) type 1. It is
CC   isolated from clone sfl1.pk0008.a4, obtained from sfl1 cDNA library,
CC   prepared from soybean immature flowers. PC4 molecules functions as a
CC   signal mediator between activators like Gal4/VP16 and general
CC   transcription factors (GTFs) in a transcription initiation complex. The
CC   expression levels of PC4 can be manipulated and the functional properties
CC   of specific transcriptional activators can be modulated. The PC4 DNA can
CC   be used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 672 BP; 218 A; 113 C; 154 G; 187 T; 0 U; 0 Other;

  Query Match            17.3%;  Score 109.4;  DB 3;  Length 672;
  Best Local Similarity  66.1%;  Pred. No. 2.5e-18;
  Matches  158;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        127 TCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCG 186
              |||||   ||   ||||  |||   |||   | |||| ||||||||||||||||| ||
Db        130 TCCGATGACGATCCCGACTCTATTACCGTTTGCGAGATTTCGAAGAACAGGAGGGTTGCC 189

Qy        187 GTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGC 246
              ||| ||| |||||| ||||   |  | || ||||| |||||||| ||||  || || |||
Db        190 GTGAGGAACTGGAAAGGCAGCATTATGGTTGACATTCGCGAGTTTTACGTCAAAGATGGC 249

Qy        247 AAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGG 306
              |||    | || ||| | ||||||||     |  | ||||||||||||||  | ||  |
Db        250 AAGCAATTGCCTGGCAGAAAAGGTATCTCTTTGACCATGGATCAGTGGAATGTGCTTCGT 309

Qy        307 GACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCT 365
              |  || |  |||  || ||  |||| || || |||||| | |     | |  ||||||
Db        310 AATCATGTTGAAGAAATTGACAAGGCAATTAATGAGAATTCTTAGGAAGTGGGCATTCT 368


RESULT 9
AAC38126
ID   AAC38126 standard; DNA; 471 BP.
XX
AC   AAC38126;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 19870.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
```

PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.

```
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match           15.2%;  Score 96;  DB 3;  Length 471;
  Best Local Similarity   66.3%;  Pred. No. 6.8e-15;
  Matches  138;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy        142 GACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAAC 201
              ||||||||||||||||  | | ||||| ||||| ||||| ||  | ||  |||  ||||||
Db        220 GACGATATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAAC 279

Qy        202 GGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGC 261
              || ||| |     | ||||| || |||||||| |  |||||||| |||||  | || |||
Db        280 GGGAAGATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGC 339

Qy        262 CGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCT 321
              ||||||||    ||     |||||||||||| |  ||  || ||| |      | | |
Db        340 AAGAAAGGTATCTCTCTAAGCGTGGATCAGTGGAACACTCTTCGGAACCACGCAGAGGAT 399

Qy        322 ATAGATGAGGCCATCAAGGAGAATGCGT 349
              || ||  ||||| ||   ||   | | |
Db        400 ATCGAGAAGGCCCTCTCTGACCTTTCTT 427


RESULT 10
AAZ50416
ID   AAZ50416 standard; cDNA; 516 BP.
XX
AC   AAZ50416;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Vernonia transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   vernonia; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone vs1n.pk013.f21; ss.
XX
OS   Vernonia mespilifolia.
XX
FH   Key             Location/Qualifiers
FT   mat_peptide     1. .333
FT                   /*tag=  a
FT                   /product= "Vernonia PC4(P15) type I transcription
FT                   coactivator"
FT                   /note= "Derived from clone vs1n.pk013.f21"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44883.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 38; 43pp; English.
XX
CC   The present sequence is the cDNA encoding vernonia Positive Cofactor 4
CC   (PC4) transcription coactivator, designated as PC4(P15) type 1. It is
CC   isolated from clone vs1n.pk013.f21, obtained from vs1n cDNA library,
CC   prepared from vernonia seeds. PC4 molecules functions as a signal
CC   mediator between activators like Gal4/VP16 and general transcription
CC   factors (GTFs) in a transcription initiation complex. The expression
CC   levels of PC4 can be manipulated and the functional properties of
CC   specific transcriptional activators can be modulated. The PC4 DNA can be
CC   used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 516 BP; 153 A; 103 C; 121 G; 126 T; 0 U; 13 Other;

  Query Match           12.4%;  Score 78.6;  DB 3;  Length 516;
  Best Local Similarity   60.6%;  Pred. No. 2.4e-10;
  Matches  120;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              || ||| ||||| ||| || |     |  | || || ||| |||||||  | ||  |||
Db        129 CGACGAAGACGACATCTTCATTTGTGACGTTTCCAAAAACCGGAGGGTTTCTGTCAGGAA 188

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              ||| | || | |||| | ||||| ||||| | ||| |||  |||  |||||||     |
Db        189 NTGGCAAGGGAGGGTCTTTGTCGATATCCGTGNGTTTTACATGAAANACGGCAANNAAAT 248

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
              || | |   | ||| ||      |  | |||||||| |||||   |||  | |   || |
Db        249 GCCTGNCAAAANAGGCATCTCATTGACCATGGATCANTGGAAAGAACTCCGTGCTCATGT 308

Qy        315 CAAAGCTATAGATGAGGC 332
              | |  |  ||  ||||
Db        309 GGATGAAANTGACAAGGC 326


RESULT 11
AAZ50418
ID   AAZ50418 standard; cDNA; 939 BP.
XX
AC   AAZ50418;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Corn transcription coactivator PC4(P15) type 2 cDNA from contig-2.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 2;
KW   corn; signal mediator; activator; Gal4/VP16; transgenic plant; contig;
```

```
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone p0118.chsbi09r;
KW   clone cpd1c.pk006.i3; clone cbn10.pk0063.h8; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             31. .453
FT                   /*tag= a
FT                   /product= "Corn PC4(P15) type 2 transcription coactivator
FT                   -2"
FT                   /note= "Derived from contig composed of clones
FT                   p0118.chsbi09r, cpd1c.pk006.i3, cbn10.pk0063.h8"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44885.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 21; Page 40-41; 43pp; English.
XX
CC   The present sequence is a cDNA encoding corn Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 2. It is from a
CC   contig composed of clones p0118.chsbi09r, cpd1c.pk006.i3, and
CC   cbn10.pk0063.h8 obtained from p0018 (corn stem tissue pooled from the 4-5
CC   internodes subtending the tassel at stages V8-V12), cpd1c (corn pooled
CC   BMS treated with chemicals related to protein kinases) and cbn10 (corn
CC   developing kernel 10 days after pollination) cDNA libraries. PC4
CC   molecules functions as a signal mediator between activators like
CC   Gal4/VP16 and general transcription factors (GTFs) in a transcription
CC   initiation complex. The expression levels of PC4 can be manipulated and
CC   the functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 939 BP; 261 A; 168 C; 275 G; 235 T; 0 U; 0 Other;

  Query Match           12.1%;  Score 76.4;  DB 3;  Length 939;
  Best Local Similarity  59.0%;  Pred. No. 1.1e-09;
  Matches  131;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              || | | |  ||  |  || |   ||   | ||||  || |||||||| | |   |
Db        228 CGACCAAGGTGACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGACTTTATCGGA 287

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              |  || |||| |    | ||   ||||||||||||||| ||||||||||||    |
Db        288 GTTTAAGGGCAGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGGCAAGGAGAT 347

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
               |||  |  |||||||||     |  |  |||| |||||| |   | |  |  |   | |
Db        348 GCCCTCCGCCAAAGGTATTAGTATGACTTTGGAGCAGTGGGAAGCATTTTGCAATGCTGT 407

Qy        315 CAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
                 || ||||| || ||||||||  ||  ||     || ||
Db        408 ACCTGCAATAGAGGATGCCATCAAAAAGCTTGAAGATTCAGA 449


RESULT 12
AAZ50417
ID   AAZ50417 standard; cDNA; 1089 BP.
XX
AC   AAZ50417;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Corn transcription coactivator PC4(P15) type 2 cDNA from contig-1.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 2;
KW   corn; signal mediator; activator; Gal4/VP16; transgenic plant; contig;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone p0014.ctuth59r;
KW   clone ceb5.pk0070.e3; clone cpi1c.pk017.j22; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             104. .655
FT                   /*tag= a
FT                   /product= "Corn PC4(P15) type 2 transcription coactivator
FT                   -1"
FT                   /note= "Derived from contig composed of clones
FT                   p0014.ctuth59r, ceb5.pk0070.e3 and cpi1c.pk017.j22"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44884.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 21; Page 39-40; 43pp; English.
XX
CC   The present sequence is a cDNA encoding corn Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 2. It is isolated
CC   from a contig composed of clones p0014.ctuth59r, ceb5.pk0070.e3 and
CC   cpi1c.pk017.j22, obtained from p0014 (corn leaf), ceb5 (corn embryo 30
CC   days after pollination) and cpi1c (corn pooled BMS treated with chemicals
CC   related to biochemical compound synthesis) cDNA libraries. PC4 molecules
CC   functions as a signal mediator between activators like Gal4/VP16 and
CC   general transcription factors (GTFs) in a transcription initiation
CC   complex. The expression levels of PC4 can be manipulated and the
CC   functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 1089 BP; 283 A; 229 C; 321 G; 255 T; 0 U; 1 Other;

  Query Match           12.0%;  Score 76;  DB 3;  Length 1089;
  Best Local Similarity  51.8%;  Pred. No. 1.5e-09;
  Matches  172;  Conservative    0;  Mismatches  160;  Indels    0;  Gaps    0;
```

```
Qy          25 GAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC 84
               ||||| |  | ||  ||     |     | | ||||   || |   |   | ||  | |
Db         320 GAGGGCGTCACTGGCAGGGAGAGCAAGGGCAAGGAACACGAAGAGGAGGACGAAGAGGAT 379

Qy          85 GAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGAC 144
               ||    | ||   ||       |      || | ||    |       |    | | | |
Db         380 GACGATGAGGAGGAGGATGAAGGTAAGGGTGGCGGGAAGAGAGAGTACGATGACCAAGGT 439

Qy         145 GATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGC 204
               ||  |  || |   ||   | ||||  || |||||||| |  |   |   |  || |||
Db         440 GACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGACTTTATCGGAGTTTAAGGGC 499

Qy         205 AAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGC 264
               | |     | ||   |||||||||||||||| ||||||||||||    | |||  |  |
Db         500 AGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGGCAAGGAGATGCCCTCCGCC 559

Qy         265 AAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATA 324
               ||||||||      |  |  |||| |||||| |   | |  |  |   | |    || |||
Db         560 AAAGGTATTAGTATGACTTTGGAGCAGTGGGAAGCATTTTGCAATGCTGTACCTGCAATA 619

Qy         325 GATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
               || || ||||||||  ||  ||     || ||
Db         620 GAGGATGCCATCAAAAAGCTTGAAGATTCAGA 651


RESULT 13
AAC32760
ID   AAC32760 standard; DNA; 1063 BP.
XX
AC   AAC32760;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 548.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
```

PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.7%; Score 74; DB 3; Length 1063;
Best Local Similarity 73.1%; Pred. No. 5e-09;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
Qy      142 GACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAAC 201
            ||||||||||||||  | | ||||| ||||| ||||| ||  | ||  |||  ||||||
Db      220 GACGATATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAAC 279

Qy      202 GGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGC 261
            || ||| |    | ||||| || |||||||| |  |||||||| |||||  | || |||
Db      280 GGGAAGATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGC 339

Qy      262 CGCAAAGGTA 271
            |||||||
Db      340 AAGAAAGGTA 349
```

RESULT 14
ACF03526
ID ACF03526 standard; cDNA; 913 BP.
XX
AC ACF03526;
XX
DT 27-OCT-2003 (revised)
DT 12-SEP-2003 (first entry)
XX
DE Brassica campestris MIP102 encoding cDNA SEQ ID NO:1.
XX
KW Brassica campestris; plant; MIP102; virus resistance;
KW plant virus transport protein binding protein; gene; ss.
XX
OS Brassica rapa.
XX
FH Key Location/Qualifiers
FT CDS 14. .511
FT /*tag= a
FT /product= "MIP102"
XX
PN WO2003022039-A1.
XX
PD 20-MAR-2003.
XX
PF 10-SEP-2001; 2001WO-JP007858.
XX
PR 10-SEP-2001; 2001WO-JP007858.
XX
PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PI Nishiguchi M, Nyunoya H, Matsushita Y;
XX
DR WPI; 2003-313169/30.
DR P-PSDB; ABR57409.
XX
PT Plant protein binding to plant virus transport protein for imparting
PT virus resistance to plants including tobacco.
XX
PS Claim 3; Fig 2; 84pp; Japanese.
XX
CC The present invention describes a method for imparting virus resistance
CC to plants in which the plant cells are transformed with a polynucleotide
CC encoding a protein (I) which binds to plant virus transport protein. Also
CC described are plants transformed by (I). The method is useful for
CC increasing resistance of plants such as tobacco, tomato, petunia, pear,
CC rice, wheat, barley, maize, soybean, oilseed rape, rose, apple, alfalfa,

```
CC   melon, lavender, onion, spinach and parsley; to viruses including
CC   tobamovirus, tobravirus, dianthovirus, alfamovirus, bromovirus,
CC   cucumovirus, comovirus, nepovirus, caulimovirus, geminivirus, potivirus
CC   and tospovirus. The present sequence encodes Brassica campestris MIP102,
CC   which is a specifically claimed plant virus transport protein binding
CC   protein from the present invention. (Updated on 27-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 913 BP; 306 A; 161 C; 224 G; 222 T; 0 U; 0 Other;

  Query Match             10.8%;  Score 68.2;  DB 8;  Length 913;
  Best Local Similarity   57.9%;  Pred. No. 1.5e-07;
  Matches  121;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              ||  |||| |||| || || |   |  | | ||| | || ||||| ||| || | | |
Db        286 CGATGACGGCGATCTCATCATTTGCAGGCTGTCGGATAAGAGGAGAGTGACGATTCAGGA 345

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              |  |  || |||    | ||   |||| | |||| ||| |||| |||||||||    ||
Db        346 GTTTAGAGGAAAGAGTTTGGTTTCCATCAGAGAGTATTACAAGAAAGACGGCAAAGAGCT 405

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
              ||       ||||| |||    |  ||   || || |||   |   | | | | |||||
Db        406 TCCTTCTTCTAAAGGAATAAGCTTAACAGACGAACAATGGTCAACGTTCAAGAAAAATAT 465

Qy        315 CAAAGCTATAGATGAGGCCATCAAGGAGA 343
                |||||| || |  ||  ||||| | |
Db        466 TCCAGCTATCGAAGCTGCTGTCAAGAAAA 494


RESULT 15
AAZ50419
ID   AAZ50419 standard; cDNA; 740 BP.
XX
AC   AAZ50419;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Soybean transcription coactivator PC4(P15) type 2 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 2;
KW   soybean; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone ses4d.pk0016.g2; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             64. .489
FT                   /*tag=  a
FT                   /product= "Soybean PC4(P15) type 2 transcription
FT                   coactivator"
FT                   /note= "Derived from clone ses4d.pk0016.g2"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44886.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 21; Page 41-42; 43pp; English.
XX
CC   The present sequence is the cDNA encoding soybean Positive Cofactor 4
CC   (PC4) transcription coactivator, designated as PC4(P15) type 2. It is
CC   isolated from clone ses4d.pk0016.g2, obtained from ses4d cDNA library,
CC   prepared from soybean embryogenic suspension 4 days after subculture. PC4
CC   molecules functions as a signal mediator between activators like
CC   Gal4/VP16 and general transcription factors (GTFs) in a transcription
CC   initiation complex. The expression levels of PC4 can be manipulated and
CC   the functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 740 BP; 249 A; 136 C; 182 G; 173 T; 0 U; 0 Other;

  Query Match             10.5%;  Score 66.2;  DB 3;  Length 740;
  Best Local Similarity   55.4%;  Pred. No. 4.7e-07;
  Matches  128;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        139 GACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGG 198
              || | |||| || || ||  |  | | ||  | || || |||||| || | | |   |
Db        271 GAAGGCGATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTGACGATTCAGGATTTC 330

Qy        199 AACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCC 258
              |  || ||     | |||  ||| || ||||  ||  | ||||| ||||||   || ||
Db        331 AGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGATGGCAAGGAACTTCCT 390

Qy        259 GGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAA 318
                |||||| |||    |  ||   || ||||||      | | | | ||| |
Db        391 ACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAGCCTTTAAGAAAAATGTGCCT 450

Qy        319 GCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTG 369
              || |||||  | ||||| ||| | |  | || |       | |||    |||
Db        451 GCCATAGAAAAAGCCATTAAGAAAATGGAGTCAAGTTGACACATGGCTTTG 501


Search completed: January 11, 2005, 20:15:08
Job time : 446 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 18:41:07 ; Search time 72 Seconds
(without alignments)
6239.147 Million cell updates/sec

Title: US-10-629-953-3
Perfect score: 632
Sequence: 1 ctcctccaagactccacaag.....taaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	182.2	28.8	289	4	US-09-313-294A-1459	Sequence 1459, Ap
	2	149.6	23.7	263	4	US-09-313-294A-1612	Sequence 1612, Ap
	3	145.6	23.0	286	4	US-09-313-294A-6916	Sequence 6916, Ap
c	4	55.4	8.8	7218	1	US-08-232-463-14	Sequence 14, Appl
	5	52.4	8.3	1074	3	US-09-248-335-67	Sequence 67, Appl
	6	45.8	7.2	954	4	US-09-252-991A-2562	Sequence 2562, Ap
	7	45.8	7.2	1953	4	US-09-252-991A-2481	Sequence 2481, Ap
c	8	45.8	7.2	2907	4	US-09-252-991A-2354	Sequence 2354, Ap
	9	45	7.1	860	3	US-09-248-335-41	Sequence 41, Appl
	10	45	7.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	11	45	7.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	12	43.4	6.9	1623	4	US-09-252-991A-15700	Sequence 15700, A
	13	43.4	6.9	2208	4	US-09-252-991A-15669	Sequence 15669, A
	14	43.2	6.8	2460	4	US-09-252-991A-2998	Sequence 2998, Ap
c	15	43.2	6.8	2856	4	US-09-252-991A-2869	Sequence 2869, Ap
	16	43.2	6.8	3387	4	US-09-252-991A-3101	Sequence 3101, Ap
c	17	42.8	6.8	675	3	US-08-998-416-723	Sequence 723, App
c	18	42	6.6	49377	1	US-08-764-233A-1	Sequence 1, Appli
c	19	42	6.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	20	42	6.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	21	41.4	6.6	1395	4	US-09-252-991A-13078	Sequence 13078, A
	22	41.4	6.6	2157	4	US-09-252-991A-12861	Sequence 12861, A
c	23	41.2	6.5	768	4	US-09-252-991A-1114	Sequence 1114, Ap
	24	41.2	6.5	924	4	US-09-252-991A-969	Sequence 969, App
	25	41.2	6.5	939	4	US-09-252-991A-925	Sequence 925, App
	26	41	6.5	1109	3	US-09-178-610-1	Sequence 1, Appli
	27	40.8	6.5	832	4	US-09-621-976-2813	Sequence 2813, Ap
	28	40.8	6.5	2943	1	US-08-042-747A-7	Sequence 7, Appli
c	29	40.6	6.4	708	4	US-09-252-991A-9083	Sequence 9083, Ap
	30	40.6	6.4	840	3	US-09-248-335-45	Sequence 45, Appl
c	31	40.6	6.4	1221	4	US-09-252-991A-9253	Sequence 9253, Ap
	32	40.6	6.4	1749	4	US-09-252-991A-8795	Sequence 8795, Ap
	33	40.6	6.4	2313	4	US-09-252-991A-8696	Sequence 8696, Ap
c	34	40.4	6.4	44377	2	US-08-804-227C-7	Sequence 7, Appli
c	35	40.4	6.4	44377	2	US-08-804-198-1	Sequence 1, Appli
	36	40.2	6.4	900	4	US-09-252-991A-9282	Sequence 9282, Ap
	37	40.2	6.4	3729	4	US-09-252-991A-9272	Sequence 9272, Ap
c	38	40	6.3	1134	4	US-09-252-991A-16141	Sequence 16141, A
	39	40	6.3	1332	4	US-09-266-965-64	Sequence 64, Appl
	40	40	6.3	2157	4	US-09-252-991A-16462	Sequence 16462, A
	41	40	6.3	4052	4	US-08-624-447-11	Sequence 11, Appl
c	42	40	6.3	53500	4	US-09-266-965-76	Sequence 76, Appl
	43	39.8	6.3	414	4	US-09-252-991A-15011	Sequence 15011, A
	44	39.8	6.3	678	4	US-09-252-991A-12248	Sequence 12248, A
c	45	39.8	6.3	969	4	US-09-252-991A-12362	Sequence 12362, A

ALIGNMENTS

RESULT 1
US-09-313-294A-1459
; Sequence 1459, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1459
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550983H1
US-09-313-294A-1459

Query Match 28.8%; Score 182.2; DB 4; Length 289;
Best Local Similarity 79.6%; Pred. No. 9.3e-44;
Matches 230; Conservative 0; Mismatches 53; Indels 6; Gaps 1;

```
Qy          41 GAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCGG 94
               | |  |||||| |||| || || ||||| |||||||| ||||||      ||||| ||||
Db           1 GGAAGATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGG 60

Qy          95 CCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCG 154
               |||||||||  |||||  |||||||  || ||||||||    |||  ||  |||| ||  |||||
Db          61 CCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120

Qy         155 TCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCG 214
               | ||| |||||||||||||||  || |||| | ||  |||  |||||| |||| |||| |||
Db         121 TAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCG 180

Qy         215 TCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC 274
               |||||  | ||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         181 TCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATAT 240

Qy         275 AGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTAT 323
                  |||| | | |||||||||||||||| ||| |||||| |||||||| ||
Db         241 CACTCCAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCCAT 289
```

RESULT 2
US-09-313-294A-1612

```
; Sequence 1612, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 1612
;   LENGTH: 263
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700551220H1
;   NAME/KEY: unsure
;   LOCATION: 25, 181
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-1612

  Query Match             23.7%;  Score 149.6;  DB 4;  Length 263;
  Best Local Similarity   78.0%;  Pred. No. 3.1e-34;
  Matches  206;  Conservative    0;  Mismatches   51;  Indels    7;  Gaps    2;

Qy          41 GAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCGG 94
               | |  |||||| |||| || || | ||| ||||||| ||||||      ||||| ||||
Db           1 GGAAGATGTGGGGGAAAGGAAAGANGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGG 60

Qy          95 CCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCG 154
               |||||||||  |||||  ||||||| || ||||||||   ||| || |||| || |||||
Db          61 CCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCG 120

Qy         155 TCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCG 214
               | ||| |||||||||||||||| || |||| | ||  ||| |||||| |||| |||| |||
Db         121 TAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCG 180

Qy         215 TCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC 274
                |||| | ||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         181 NCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA- 239

Qy         275 AGCTCCCAATGGATCAGTGGAAGA 298
                  ||| | | ||||||||||||||
Db         240 TCATCCAATTAGATCAGTGGAAGA 263


RESULT 3
US-09-313-294A-6916
; Sequence 6916, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 6916
;   LENGTH: 286
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700380835H1
;   NAME/KEY: unsure
;   LOCATION: 54
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6916

  Query Match             23.0%;  Score 145.6;  DB 4;  Length 286;
  Best Local Similarity   78.4%;  Pred. No. 4.9e-33;
  Matches  189;  Conservative    0;  Mismatches   45;  Indels    7;  Gaps    1;

Qy          40 AGAAGAATGTGGCGGAAGGGGAACAAGCGGTTC-------GGCGGCGGCGGCGAGCCGGC 92
               || |  |||| | |||| || || ||||| |||       |||||| |||| ||||| ||
Db          44 AGGAAGATGTNGGGGAAAGGAAAGAAGCGTTTCGGGCGGTGGCGGCAGCGGTGAGCCCGC 103

Qy          93 GGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGT 152
               ||||||||||| |||||  ||||||| || ||||||||    ||| || |||| || |||
Db         104 GGCCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGT 163

Qy         153 CGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGT 212
               ||| ||| |||||||||||||||| || |||| | ||  ||| |||||| |||| |||| |
Db         164 CGTAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTT 223

Qy         213 CGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTAT 272
               ||||||| | |||||||||||||   |||||||||||||| ||||||  ||||||||||||
Db         224 CGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTAT 283

Qy         273 A 273
               |
Db         284 A 284


RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
```

```
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             8.8%;  Score 55.4;  DB 1;  Length 7218;
  Best Local Similarity   6.7%;  Pred. No. 7.5e-06;
  Matches   29;  Conservative  223;  Mismatches  179;  Indels    0;  Gaps    0;

Qy         11 ACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGT 70
              ||    : :::   :::::::::::: : :::::::: : :: :::::::::: ::: ::
Db       1436 ACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1377

Qy         71 TCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCG 130
                :: :: :: :: ::: :: ::  ::: :  : :  :  ::::: :: :::      :
Db       1376 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1317

Qy        131 AGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGC 190
              :::: :  :: :: :: : :  :  :  :  ::: :  :::::: :::::::: :: :: :
Db       1316 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1257

Qy        191 GGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGA 250
              :::   :::: :: ::::  :  :  :: :   : :::    : :::::::: :: ::::
Db       1256 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1197

Qy        251 CCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACA 310
                  ::  : ::::: : : ::     :: :::  :: :::::: :  :::::: :
Db       1196 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1137

Qy        311 ATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTGT 370
              : :  :::  : ::: :::  :  :::::::: : : ::  ::::   :      :
Db       1136 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1077

Qy        371 GATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATATTGCACTTTTG 430
              :: : ::: :  |  |  || | ||   ||      |  |   | | || | |  |
Db       1076 RRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGGAATTAATTCTGTGAGC 1017

Qy        431 GTATGGTAATC 441
              |||||| || |
Db       1016 GTATGGCAAAC 1006


RESULT 5
US-09-248-335-67
; Sequence 67, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 67
;   LENGTH: 1074
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-67

  Query Match             8.3%;  Score 52.4;  DB 3;  Length 1074;
  Best Local Similarity   52.3%;  Pred. No. 2e-05;
  Matches  116;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy         71 TCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCG 130
              || | | | ||| || ||       |||| ||| |||  | |  |||  | |  |  |
Db         91 TCAGGGCCCGCGTCGCGCTAAACCTCAAGGGCCTGGCCTACCGATACGTGGAGGACAACC 150

Qy        131 AGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGC 190
              |  | | | |   |||  || || |||||    |   |   | || ||| |||| ||||||
Db        151 TGGACAGCAAGAGCGAGCTCCTCCTCGCCTCCAACCCCGTGCACGGGAAGGTGCCGGTGC 210

Qy        191 GGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGA 250
              |   ||||| |  ||||  |||   |||  |  ||   ||  ||| |   ||  | ||
Db        211 TCCTCCACGACGGCAGGCCCGTCTGCGAGTCCCGGGTCATCGTGGAGTATATCGACGAGG 270

Qy        251 CCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGT 292
              || |||| | | ||          |  |||||    || | ||
Db        271 CCTTCCCGGCCAGCGGCCCCTGCCTCCTCCCCGCCGACCCGT 312


RESULT 6
US-09-252-991A-2562
; Sequence 2562, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 2562
;   LENGTH: 954
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2562

  Query Match             7.2%;  Score 45.8;  DB 4;  Length 954;
  Best Local Similarity   53.7%;  Pred. No. 0.0016;
  Matches   95;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy         71 TCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCG 130
              ||||||  |   |  | ||  |||||||||       | |  || | ||| |||  ||   |
Db        299 TCGGCGACCTGGCGGTGCGCGGCCAAGGCGGTGGGCTACGAGAACGCCGGCACCGTGG 358

Qy        131 AGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGC 190
              ||  |     || |||   ||  ||   |   ||   || |||||   ||||   ||||
Db        359 AGTTCCTGCTCGCCGACGGCGAGGTGTACTTCATGGAGATGAACACCCGGGTCCAGGTGG 418

Qy        191 GGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCA 247
              |  |   | |  | |||   ||  || ||||  || | ||  |||| ||  | |||
Db        419 AGCACACCATCACCGAGGAAATCACCGGCATCGACGTGGTCCGCGAGCAGATCCGCA 475


RESULT 7
US-09-252-991A-2481
; Sequence 2481, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 2481
;   LENGTH: 1953
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2481

  Query Match             7.2%;  Score 45.8;  DB 4;  Length 1953;
  Best Local Similarity   53.7%;  Pred. No. 0.0024;
  Matches   95;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;
```

```
Qy          71 TCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCG 130
               |||||| |    |  | ||  |||||||||       | |  || | ||| |||  ||   |
Db         894 TCGGCGACCTGGCGGTGCGCGCGGCCAAGGCGGTGGGCTACGAGAACGCCGGCACCGTGG 953

Qy         131 AGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGC 190
               ||  |     || |||   ||  ||   |   ||   || |||||   ||||   ||||
Db         954 AGTTCCTGCTCGCCGACGGCGAGGTGTACTTCATGGAGATGAACACCCGGGTCCAGGTGG 1013

Qy         191 GGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCA 247
               |  |    | |  | |||   ||  || ||||  || | ||  |||| ||  | |||
Db        1014 AGCACACCATCACCGAGGAAATCACCGGCATCGACGTGGTCCGCGAGCAGATCCGCA 1070


RESULT 8
US-09-252-991A-2354/c
; Sequence 2354, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 2354
;   LENGTH: 2907
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2354

  Query Match             7.2%;  Score 45.8;  DB 4;  Length 2907;
  Best Local Similarity   53.7%;  Pred. No. 0.003;
  Matches   95;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy          71 TCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCG 130
               |||||| |    |  | ||  |||||||||       | |  || | ||| |||  ||   |
Db        2612 TCGGCGACCTGGCGGTGCGCGCGGCCAAGGCGGTGGGCTACGAGAACGCCGGCACCGTGG 2553

Qy         131 AGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGC 190
               ||  |     || |||   ||  ||   |   ||   || |||||   ||||   ||||
Db        2552 AGTTCCTGCTCGCCGACGGCGAGGTGTACTTCATGGAGATGAACACCCGGGTCCAGGTGG 2493

Qy         191 GGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCA 247
               |  |    | |  | |||   ||  || ||||  || | ||  |||| ||  | |||
Db        2492 AGCACACCATCACCGAGGAAATCACCGGCATCGACGTGGTCCGCGAGCAGATCCGCA 2436


RESULT 9
US-09-248-335-41
; Sequence 41, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 41
;   LENGTH: 860
;   TYPE: DNA
;   ORGANISM: maize
US-09-248-335-41

  Query Match             7.1%;  Score 45;  DB 3;  Length 860;
  Best Local Similarity   50.2%;  Pred. No. 0.0026;
  Matches  111;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy          69 GTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTC 128
               |||  |||    ||   | || |    ||||  |||    |  || | ||  || ||
Db          95 GTTGCGCGTGCGCATGGCGCTGAGCCTCAAGGGCCTGAGCTACGAGTACATCGAGCAGGA 154

Qy         129 CGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGT 188
               |  |  |  | | | |||  || || ||  |   |   |   | ||| || |||| | ||
Db         155 CCTGTTCCACAAGGGCGAGCTCCTCCTCAGCTCAAACCCCGTGCACAAGAAGGTGCCCGT 214

Qy         189 GCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAA 248
               ||  | |    |||||||  | ||  |||   || ||   ||   ||| | | ||   |
Db         215 GCTCATCCACCACGGCAAGCCCATCTGCGAGTCCCTCGCCGTCGTGGAGTACGTCGATGA 274

Qy         249 GACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATC 289
               |  |   ||||||  |   |  |     |||||    || |
Db         275 GGTCTGGCCCGGCGCCGCCGCCACCATCCTCCCCGCCGACC 315


RESULT 10
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             7.1%;  Score 45;  DB 3;  Length 4403765;
  Best Local Similarity   50.7%;  Pred. No. 0.28;
  Matches  108;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy          58 GGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGAC 117
               |||  |  | | | ||||||   |||||   | |  |      |  | ||||  |||| |
Db     925447 GGGGCCGGGGGTTCCGGCGGTTCCGGCGCAGCTGGTGGGGCCGGTGGCGCCGGTGGGGCC 925506

Qy         118 GACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGG 177
               | |||||  | ||  |||| || |  ||      ||| ||       |||||  |  ||
Db     925507 GGCGGGCTGTTCGCCAGCGGCGGCAGCGGCGGGTTCGGCGGGTTCGCATCGACGGGCACC 925566

Qy         178 AGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAG 237
               | | ||| |   | ||| |     |      |   ||| || | ||| | ||  ||
Db     925567 GGTGGGGCCGGCGGCACCGGTGGGGCTGGTGGGTTGTTCGCCAGCGGCGGGGTCGGCGGT 925626

Qy         238 AAGGACGGCAAGACCCTCCCCGGCCGCAAAGGT 270
               |  | |||     ||    ||||| | |  |||
Db     925627 ACTGGCGGGGGAGCCGGGTCCGGCGGTACCGGT 925659


RESULT 11
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
```

```
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             7.1%;  Score 45;  DB 3;  Length 4411529;
  Best Local Similarity   50.7%;  Pred. No. 0.28;
  Matches  108;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy         58 GGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGAC 117
              |||  |  | | | ||||||   |||||   | |  |       |  | ||||  |||| |
Db     925548 GGGGCCGGGGGTTCCGGCGGTTCCGGCGCAGCTGGTGGGGCCGGTGGCGCCGGTGGGGCC 925607

Qy        118 GACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGG 177
              | |||||  | ||  |||| || |  ||      ||| ||        |||||  |  ||
Db     925608 GGCGGGCTGTTCGCCAGCGGCGGCAGCGGCGGGTTCGGCGGGTTCGCATCGACGGGCACC 925667

Qy        178 AGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAG 237
              | | ||| |   | ||| |    |     |   | ||| || | ||| | ||  ||
Db     925668 GGTGGGGCCGGCGGCACCGGTGGGGCTGGTGGGTTGTTCGCCAGCGGCGGGGTCGGCGGT 925727

Qy        238 AAGGACGGCAAGACCCTCCCCGGCCGCAAAGGT 270
              |  | |||      ||    ||||| | |  |||
Db     925728 ACTGGCGGGGGAGCCGGGTCCGGCGGTACCGGT 925760


RESULT 12
US-09-252-991A-15700
; Sequence 15700, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 15700
;   LENGTH: 1623
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15700

  Query Match             6.9%;  Score 43.4;  DB 4;  Length 1623;
  Best Local Similarity   51.9%;  Pred. No. 0.011;
  Matches   98;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy         78 CGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGC 137
              |||||  | ||||    | |  ||  |   ||  ||| |||| || |    |||||| |||
Db       1005 CGGCGCCAAGCCCAGCGTCCTGCCGCTGTCCATCGGCGACGGCGAGTTGGCCGAGACCGC 1064

Qy        138 CGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTG 197
              |||   ||   ||  |  || | |||| |       || ||  |  ||  |  |||| |
Db       1065 CGATACCGTGGTCCCCACCGGCGAGATCTTCCGCTACTGGCTGCAGGGCGGACGGATCGA 1124

Qy        198 GAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCC 257
              ||||    ||| ||  | |      |||   || ||  ||   |  ||  ||| |
Db       1125 CGTCGGCTTCCTCGGCGCGGCCCAGGTCGACCGCTTCGGCAACATCAACACCACCGTGAT 1184

Qy        258 CGGCCGCAA 266
              ||||  | |
Db       1185 CGGCGACTA 1193


RESULT 13
US-09-252-991A-15669
; Sequence 15669, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 15669
;   LENGTH: 2208
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15669

  Query Match             6.9%;  Score 43.4;  DB 4;  Length 2208;
  Best Local Similarity   51.9%;  Pred. No. 0.013;
  Matches   98;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy         78 CGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGC 137
              |||||  | ||||    | |  ||  |   ||  ||| |||| || |    |||||| |||
Db        388 CGGCGCCAAGCCCAGCGTCCTGCCGCTGTCCATCGGCGACGGCGAGTTGGCCGAGACCGC 447

Qy        138 CGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTG 197
              |||   ||   ||  |  || | |||| |       || ||  |  ||  |  |||| |
Db        448 CGATACCGTGGTCCCCACCGGCGAGATCTTCCGCTACTGGCTGCAGGGCGGACGGATCGA 507

Qy        198 GAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCC 257
              ||||    ||| ||  | |      |||   || ||  ||   |  ||  ||| |
Db        508 CGTCGGCTTCCTCGGCGCGGCCCAGGTCGACCGCTTCGGCAACATCAACACCACCGTGAT 567

Qy        258 CGGCCGCAA 266
              ||||  | |
Db        568 CGGCGACTA 576


RESULT 14
US-09-252-991A-2998
; Sequence 2998, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 2998
;   LENGTH: 2460
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
```

US-09-252-991A-2998

Query Match 6.8%; Score 43.2; DB 4; Length 2460;
Best Local Similarity 50.5%; Pred. No. 0.016;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy         86 AGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACG 145
              | |||| |||| ||       ||| | ||  |      |  | ||    ||||   | |||
Db       1902 ACCCGGTGGCCGAGGCGATGACCGGCTGGACCTTGCAGGACGCCATCGGCGCGCCCCACG 1961

Qy        146 ATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCA 205
              |  | ||| ||  |  | |    |   |||   | | |||  ||   | |  |    |
Db       1962 AACTGGTCTTCAACGTGCTCGACAGCGACACCGGCGAGGCCCTGGAAAGCCCGGTGCGAC 2021

Qy        206 AGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCA 265
                | |     ||| ||| | |||||||  | ||||||||   || ||   | ||||||
Db       2022 GCTGCCTGCAGAACACCCGGGTGTTCTACCTGGAGGACGGCGCCACGCTGATCAGCCGCA 2081

Qy        266 AAGGTATACAGCTCCCAATGGATCAGTG 293
              | ||    |||| |   ||  |  | ||
Db       2082 AGGGCGAGCAGCACGATATCCAGAATTG 2109
```

RESULT 15
US-09-252-991A-2869/c
; Sequence 2869, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2869
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2869

Query Match 6.8%; Score 43.2; DB 4; Length 2856;
Best Local Similarity 50.5%; Pred. No. 0.017;
Matches 105; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy         86 AGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACG 145
              | |||| |||| ||       ||| | ||  |      |  | ||    ||||   | |||
Db       1625 ACCCGGTGGCCGAGGCGATGACCGGCTGGACCTTGCAGGACGCCATCGGCGCGCCCCACG 1566

Qy        146 ATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCA 205
              |  | ||| ||  |  | |    |   |||   | | |||  ||   | |  |    |
Db       1565 AACTGGTCTTCAACGTGCTCGACAGCGACACCGGCGAGGCCCTGGAAAGCCCGGTGCGAC 1506

Qy        206 AGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCA 265
                | |     ||| ||| | |||||||  | ||||||||   || ||   | ||||||
Db       1505 GCTGCCTGCAGAACACCCGGGTGTTCTACCTGGAGGACGGCGCCACGCTGATCAGCCGCA 1446

Qy        266 AAGGTATACAGCTCCCAATGGATCAGTG 293
              | ||    |||| |   ||  |  | ||
Db       1445 AGGGCGAGCAGCACGATATCCAGAATTG 1418
```

Search completed: January 11, 2005, 21:22:20
Job time : 96 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 21:06:03 ; Search time 483 Seconds
(without alignments)
7507.856 Million cell updates/sec

Title: US-10-629-953-3
Perfect score: 632
Sequence: 1 ctcctccaagactccacaag.....taaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4293498 seqs, 2868903791 residues

Total number of hits satisfying chosen parameters: 8586996

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	610.8	96.6	1754	17	US-10-437-963-83372	Sequence 83372, A
2	208	32.9	660	16	US-10-425-114-4500	Sequence 4500, Ap
3	204.8	32.4	757	18	US-10-425-115-74342	Sequence 74342, A
4	146.6	23.2	310	9	US-09-294-093B-67	Sequence 67, Appl
5	109.4	17.3	663	16	US-10-424-599-38767	Sequence 38767, A
6	102.6	16.2	917	16	US-10-424-599-38768	Sequence 38768, A
7	76	12.0	933	16	US-10-425-114-24336	Sequence 24336, A
8	76	12.0	935	16	US-10-425-114-33158	Sequence 33158, A
9	76	12.0	1225	18	US-10-425-115-154421	Sequence 154421,
10	74.8	11.8	658	16	US-10-425-114-21115	Sequence 21115, A
11	74.8	11.8	658	18	US-10-425-115-20439	Sequence 20439, A
12	73.2	11.6	471	17	US-10-767-701-29085	Sequence 29085, A
c 13	67.4	10.7	497	16	US-10-424-599-127795	Sequence 127795,
14	66.2	10.5	794	16	US-10-425-114-20666	Sequence 20666, A
15	66.2	10.5	911	16	US-10-424-599-4356	Sequence 4356, Ap
c 16	64.4	10.2	1030	17	US-10-437-963-59601	Sequence 59601, A
17	63.2	10.0	409	18	US-10-425-115-157838	Sequence 157838,
18	62.6	9.9	1479	16	US-10-424-599-9157	Sequence 9157, Ap
c 19	61	9.7	496	10	US-09-770-961-520	Sequence 520, App
20	52	8.2	892	18	US-10-425-115-83949	Sequence 83949, A
21	50.8	8.0	1168	18	US-10-425-115-150589	Sequence 150589,
c 22	50.6	8.0	1887	17	US-10-437-963-75128	Sequence 75128, A
23	49.4	7.8	1651	17	US-10-767-701-13956	Sequence 13956, A
24	49.2	7.8	900	16	US-10-425-114-22315	Sequence 22315, A
25	49.2	7.8	1015	16	US-10-425-114-25767	Sequence 25767, A
26	48.8	7.7	647	16	US-10-425-114-16885	Sequence 16885, A
27	48.8	7.7	2209	17	US-10-437-963-69229	Sequence 69229, A
c 28	48	7.6	1996	17	US-10-437-963-46655	Sequence 46655, A
29	47.8	7.6	542	17	US-10-767-701-3376	Sequence 3376, Ap
30	47.8	7.6	693	17	US-10-437-963-68177	Sequence 68177, A
31	47.6	7.5	597	16	US-10-260-238-5561	Sequence 5561, Ap
32	47.6	7.5	1071	17	US-10-437-963-94451	Sequence 94451, A
33	47.6	7.5	1397	16	US-10-425-114-19393	Sequence 19393, A
34	47.6	7.5	1755	16	US-10-425-114-904	Sequence 904, App
35	47.6	7.5	1755	18	US-10-425-115-100553	Sequence 100553,
36	47.6	7.5	1810	16	US-10-425-114-31485	Sequence 31485, A
37	47.6	7.5	2365	18	US-10-425-115-100551	Sequence 100551,
c 38	47.4	7.5	2256646	17	US-10-470-565-1	Sequence 1, Appli
39	47.2	7.5	1509	10	US-09-967-464-68	Sequence 68, Appl
40	46.8	7.4	1848	16	US-10-293-913A-1	Sequence 1, Appli
41	46.8	7.4	1848	16	US-10-293-913A-3	Sequence 3, Appli
c 42	46.8	7.4	3123	18	US-10-425-115-85044	Sequence 85044, A
43	46.6	7.4	1497	16	US-10-425-114-24433	Sequence 24433, A
44	46.6	7.4	1535	18	US-10-425-115-67209	Sequence 67209, A
45	46.4	7.3	1509	10	US-09-967-464-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-10-437-963-83372
; Sequence 83372, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83372
; LENGTH: 1754
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8270C.1
US-10-437-963-83372

Query Match 96.6%; Score 610.8; DB 17; Length 1754;
Best Local Similarity 99.7%; Pred. No. 1.7e-179;
Matches 612; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 1137
```

```
Qy          61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1138 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 1197

Qy         121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1198 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 1257

Qy         181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1258 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 1317

Qy         241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1318 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGATA 1377

Qy         301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1378 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 1437

Qy         361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1438 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 1497

Qy         421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480
               ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        1498 TGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 1557

Qy         481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1558 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 1617

Qy         541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1618 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 1677

Qy         601 GGTTTCTGCGTGTA 614
               ||||||||||||||
Db        1678 GGTTTCTGCGTGTA 1691
```

RESULT 2
US-10-425-114-4500
; Sequence 4500, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4500
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700380835_FLI
US-10-425-114-4500

Query Match 32.9%; Score 208; DB 16; Length 660;
Best Local Similarity 79.8%; Pred. No. 5e-54;
Matches 260; Conservative 0; Mismatches 60; Indels 6; Gaps 1;

```
Qy          40 AGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC------GAGCCGGCG 93
               || |  |||||| |||| || || ||||| |||||||| ||||||      ||||| |||
Db          45 AGGAAGATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCG 104

Qy          94 GCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTC 153
               ||||||||||  |||||  ||||||| || ||||||||    ||| || |||| || ||||
Db         105 GCCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTC 164

Qy         154 GTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTC 213
               || ||| ||||||||||||||| || |||| | ||  ||| |||||| |||| |||| ||
Db         165 GTAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTC 224

Qy         214 GTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATA 273
               ||||||  | ||||||||||||    |||||||||||||| ||||||  ||||||||||||
Db         225 GTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA 284

Qy         274 CAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCC 333
               ||||  | | |||||||||||||||| ||| |||||| |||||||| ||| |||||||
Db         285 TCACTCCAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCCATAAATGAGGCA 344

Qy         334 ATCAAGGAGAATGCGTGATCGGAGCC 359
               ||| |||| ||| | |||| ||||||
Db         345 ATCGAGGAAAATACATGATGGGAGCC 370

RESULT 3
US-10-425-115-74342
; Sequence 74342, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 74342
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167799C.1
US-10-425-115-74342

Query Match 32.4%; Score 204.8; DB 18; Length 757;
Best Local Similarity 74.6%; Pred. No. 5.3e-53;
Matches 256; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```
Qy   17 CAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCG 76
               | |||| || || ||| || | | | | ||| | | ||||
Db   49 CCAGTCGAGTCTCAGCAAGGAAGATGTGGGGGAGAGGAAAGAAGCGTTTCGGCGGTGGCG 108

Qy   77 GCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCG 136
               || |||| ||||| |||||||||||| ||||| ||||||| || |||||||| ||
Db   109 GCAGCGGTGAGCCCGCGGCCAAGCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCG 168

Qy   137 CCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCT 196
               | || |||| || |||||| ||| ||||||||||||||| || |||| | || ||| ||
Db   169 CTGAAGACGGTACCGTCGTAGCCGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCT 228

Qy   197 GGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCC 256
               |||| |||| |||| |||||||| | |||||||||||| |||||||||||||| ||||
Db   229 GGAAAGGCAGGGTCTTCGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCC 288

Qy   257 CCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCA 316
               || |||||||||||| |||| | | |||||||||||||||| ||| |||||| ||||
Db   289 CCACCCGCAAAGGTATATCACTCCAATTAGATCAGTGGAAGATATTGAAGGACAACATCA 348

Qy   317 AAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCC 359
               |||| ||| |||||||||| |||| ||| | |||| | |||
Db   349 AAGCCATAAATGAGGCCATCGAGGAAAATACATGATGGTTGCC 391
```
```

```
RESULT 4
US-09-294-093B-67
; Sequence 67, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
;  APPLICANT:      Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:       PERL Program
; SEQ ID NO 67
;   LENGTH: 310
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700281907H2
;   NAME/KEY: unsure
;   LOCATION: 2, 30, 295
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-67

  Query Match            23.2%;  Score 146.6;  DB 9;  Length 310;
  Best Local Similarity  71.2%;  Pred. No. 4.9e-35;
  Matches  220;  Conservative    0;  Mismatches   86;  Indels    3;  Gaps    2;

Qy         17 CAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCG 76
              | |||| ||    || |||   ||       | || | |||  |       |    ||||
Db          3 CCAGTCGAGTCTCAGCAAGGAAGATGTNGGGGAGGGGAAAGAAGCGTTTCGGCGGTGGCG 62

Qy         77 GCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCC-GGGGACGACGGGCCCTCCGAGAGC 135
              || |||||||||| ||||||||||||||  |||||  |||||||| || ||||||||    |
Db         63 GCAGCGGCGAGCCCGCGGCCAAGCGCCAGGCCGCGAGGGGACGATGGCCCCTCCGAATCC 122

Qy        136 GCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACC 195
              || || |||| || |||||| ||| ||||||||||||||| || |||  | ||  ||| |
Db        123 GCTGAAGACGGTACCGTCGTAGCCGAGATATCGAAGAACAAGAAGGTATCCGTTAGGAGC 182

Qy        196 TGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTC 255
              ||||| |||| |||| |||||||| | ||||||||||   ||||||||||||| |||
Db        183 TGGAAAGGCAGGGTCTTCGTCGACTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTC 242

Qy        256 CCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATC 315
              |||  ||||||||     ||  | |||||       ||||||||| ||| |||| | |||
Db        243 CCCACCCGCAAAGTATATCA--TTCCAATTAGATCATGGAAGATATTGAAGGACNACATC 300

Qy        316 AAAGCTATA 324
              ||||| |||
Db        301 AAAGCCATA 309


RESULT 5
US-10-424-599-38767
; Sequence 38767, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 38767
;   LENGTH: 663
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_135006C.1
US-10-424-599-38767

  Query Match            17.3%;  Score 109.4;  DB 16;  Length 663;
  Best Local Similarity  66.1%;  Pred. No. 2.9e-23;
  Matches  158;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        127 TCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCG 186
              |||||   ||    ||||  |||   |||   | |||| ||||||||||||||||| ||
Db        259 TCCGATGACGATCCCGACTCTATTACCGTTTGCGAGATTTCGAAGAACAGGAGGGTTGCC 318

Qy        187 GTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGC 246
              ||| ||| |||||| ||||   |  | || ||||| |||||||| ||||  || || |||
Db        319 GTGAGGAACTGGAAAGGCAGCATTATGGTTGACATTCGCGAGTTTTACGTCAAAGATGGC 378

Qy        247 AAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGG 306
              |||     | || ||| | ||||||||      | | ||||||||||||||  | ||  |
Db        379 AAGCAATTGCCTGGCAGAAAAGGTATCTCTTTGACCATGGATCAGTGGAATGTGCTTCGT 438

Qy        307 GACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCT 365
               |  || |  |||  || ||  |||| || || |||||| | |     | |  ||||||
Db        439 AATCATGTTGAAGAAATTGACAAGGCAATTAATGAGAATTCTTAGGAAGTGGGCATTCT 497


RESULT 6
US-10-424-599-38768
; Sequence 38768, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 38768
;   LENGTH: 917
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_135007C.1
US-10-424-599-38768

  Query Match            16.2%;  Score 102.6;  DB 16;  Length 917;
  Best Local Similarity  62.1%;  Pred. No. 4.5e-21;
  Matches  162;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy        127 TCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCG 186
              |||||   ||    ||||  | |   | |   | || ||||||||||||||||| ||
Db        341 TCCGACGACGATCCCGACTCTGTTACCATTTGCGAAATTTCGAAGAACAGGAGGGTTGCC 400

Qy        187 GTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGC 246
              ||| ||| |||||| ||||   |  | || ||||| |||||||| ||||  || || |||
Db        401 GTGAGGAACTGGAAAGGCAGCATTATGGTTGACATTCGCGAGTTTTACGTCAAAGATGGC 460

Qy        247 AAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGG 306
              |||     | || ||| | ||||||||   | |  | |||||||||||||||  | ||  |
Db        461 AAGCAATTGCCTGGCAGGAAAGGTATCTCGTTGACCATGGATCAGTGGAATGTGCTTCGT 520

Qy        307 GACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTC 366
               |  || |  |||  || ||  ||||  | || |||||| | |     | |  |||
Db        521 AATCATGTTGAAGAAATTGACAAGGCAGTTAATGAGAATTCTTAGGAAGTGGGCATCTAG 580

Qy        367 TTGTGATGCAAGTAGACTAAG 387
```

```
Db        581 TCAATTTAATAGTTTACTATG 601


RESULT 7
US-10-425-114-24336
; Sequence 24336, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24336
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3608-031-G8_FLI
US-10-425-114-24336

  Query Match             12.0%;  Score 76;  DB 16;  Length 933;
  Best Local Similarity   51.8%;  Pred. No. 9.1e-13;
  Matches  172;  Conservative    0;  Mismatches  160;  Indels    0;  Gaps    0;

Qy          25 GAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC 84
               |||||| |  | ||  ||     |    | | ||||   || |   |   | ||  | |
Db         264 GAGGGCGTCACTGGCAGGGAGAGCAAGGGCAAGGAACACGAAGAGGAGGACGAAGAGGAT 323

Qy          85 GAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGAC 144
               ||    | ||   ||       |     || | ||    |        |    | | | |
Db         324 GACGATGAGGAGGAGGATGAAGGTAAGGGTGGCGGGAAGAGAGAGTACGATGACCAAGGT 383

Qy         145 GATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGC 204
               ||  |  || |   ||   | ||||  || |||||||||| |  |   |   |  || |||
Db         384 GACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGACTTTATCGGAGTTTAAGGGC 443

Qy         205 AAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGC 264
               | |    | ||   ||||||||||||||| ||||||||||||    | |||  |  |
Db         444 AGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGGCAAGGAGATGCCCTCCGCC 503

Qy         265 AAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATA 324
               ||||||||      |  |  |||| ||||||  |   | |  |  |   | |    || |||
Db         504 AAAGGTATCAGTATGACTTTGGAGCAGTGGGAAGCATTTTGCAATGCTGTACCTGCAATA 563

Qy         325 GATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
               || ||.||||||||  ||  ||     || ||
Db         564 GAGGATGCCATCAAAAAGCTTGAAGATTCAGA 595


RESULT 8
US-10-425-114-33158
; Sequence 33158, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33158
;   LENGTH: 935
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLMO17069H06_FLI
US-10-425-114-33158

  Query Match             12.0%;  Score 76;  DB 16;  Length 935;
  Best Local Similarity   51.8%;  Pred. No. 9.1e-13;
  Matches  172;  Conservative    0;  Mismatches  160;  Indels    0;  Gaps    0;

Qy          25 GAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC 84
               ||||| |  | ||  ||     |    | | ||||   || |   |   | ||  | |
Db         282 GAGGGCGTCACTGGCAGGGAGAGCAAGGGCAAGGAACACGAAGAGGAGGACGAAGAGGAT 341

Qy          85 GAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGAC 144
               ||    | ||   ||       |     || | ||    |        |    | | | |
Db         342 GACGATGAGGAGGAGGATGAAGGTAAGGGTGGCGGGAAGAGAGAGTACGATGACCAAGGT 401

Qy         145 GATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGC 204
               ||  |  || |   ||   | ||||  || |||||||||| |  |   |   |  || |||
Db         402 GACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGACTTTATCGGAGTTTAAGGGC 461

Qy         205 AAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGC 264
               | |    | ||   ||||||||||||||| ||||||||||||    | |||  |  |
Db         462 AGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGGCAAGGAGATGCCCTCCGCC 521

Qy         265 AAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATA 324
               ||||||||      |  |  |||| ||||||  |   | |  |  |   | |    || |||
Db         522 AAAGGTATCAGTATGACTTTGGAGCAGTGGGAAGCATTTTGCAATGCTGTACCTGCAATA 581

Qy         325 GATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
               || ||.||||||||  ||  ||     || ||
Db         582 GAGGATGCCATCAAAAAGCTTGAAGATTCAGA 613


RESULT 9
US-10-425-115-154421
; Sequence 154421, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 154421
;   LENGTH: 1225
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_72412C.1
US-10-425-115-154421

  Query Match             12.0%;  Score 76;  DB 18;  Length 1225;
  Best Local Similarity   51.8%;  Pred. No. 1e-12;
  Matches  172;  Conservative    0;  Mismatches  160;  Indels    0;  Gaps    0;

Qy          25 GAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGC 84
               ||||| |  | ||  ||     |    | | ||||   || |   |   | ||  | |
Db         343 GAGGGCGTCACTGGCAGGGAGAGCAAGGGCAAGGAACACGAAGAGGAGGACGAAGAGGAT 402

Qy          85 GAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGAC 144
               ||    | ||   ||       |     || | ||    |        |    | | | |
Db         403 GACGATGAGGAGGAGGATGAAGGTAAGGGTGGCGGGAAGAGAGAGTACGATGACCAAGGT 462
```

```
Qy        145 GATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGC 204
              ||  |  || |   ||   | ||||  || |||||||||| |  |   |   |  || |||
Db        463 GACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGACTTTATCGGAGTTTAAGGGC 522

Qy        205 AAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGC 264
              | |    | ||   ||||||||||||||||| ||||||||||||    | |||  |  |
Db        523 AGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGGCAAGGAGATGCCCTCCGCC 582

Qy        265 AAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATA 324
              |||||||     |  |  |||| |||||| |   | |  |  |   | |    || |||
Db        583 AAAGGTATCAGTATGACTTTGGAGCAGTGGGAAGCATTTTGCAATGCTGTACCTGCAATA 642

Qy        325 GATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
              || || ||||||||  ||  ||     || ||
Db        643 GAGGATGCCATCAAAAAGCTTGAAGATTCAGA 674


RESULT 10
US-10-425-114-21115
; Sequence 21115, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 21115
;   LENGTH: 658
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3279-041-H1_FLI
US-10-425-114-21115

  Query Match             11.8%;  Score 74.8;  DB 16;  Length 658;
  Best Local Similarity   58.6%;  Pred. No. 1.8e-12;
  Matches  130;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              || | | |  ||  |  |  |   ||   | ||||  || |||||||||| |  ||  |
Db        229 CGACCAAGGGGACCTTATTCTGTGCCGCTTGTCGAGCAAGAGGAGGGTGACTCTGTCGGA 288

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              |  || |||| |    | ||   |||||| |||||||||| |||||||||    |||    |
Db        289 GTTCAAGGGCAGGTCGCTGGTGTCCATCCGAGAGTTCTACGTGAAGGACGGAAAGGAGAT 348

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
              |||  |  |||||||||     |  ||||||| ||||||  |   |  |  |   | |
Db        349 GCCCTCCGCCAAAGGTATTAGTATGACAATGGAGCAGTGGGAAGCATTTTGCAATGCTGT 408

Qy        315 CAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
                  || ||||| |  ||||||||  ||  ||     || ||
Db        409 ACCTGCGATAGAGGCTGCCATCAAAAAGCTTGAAGATTCAGA 450


RESULT 11
US-10-425-115-20439
; Sequence 20439, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 20439
;   LENGTH: 658
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_118642C.1
US-10-425-115-20439

  Query Match             11.8%;  Score 74.8;  DB 18;  Length 658;
  Best Local Similarity   58.6%;  Pred. No. 1.8e-12;
  Matches  130;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              || | | |  ||  |  |  |   ||   | ||||  || |||||||||| |  ||  |
Db        229 CGACCAAGGGGACCTTATTCTGTGCCGCTTGTCGAGCAAGAGGAGGGTGACTCTGTCGGA 288

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              |  || |||| |    | ||   |||||| |||||||||| |||||||||    |||    |
Db        289 GTTCAAGGGCAGGTCGCTGGTGTCCATCCGAGAGTTCTACGTGAAGGACGGAAAGGAGAT 348

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
              |||  |  |||||||||     |  ||||||| ||||||  |   |  |  |   | |
Db        349 GCCCTCCGCCAAAGGTATTAGTATGACAATGGAGCAGTGGGAAGCATTTTGCAATGCTGT 408

Qy        315 CAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
                  || ||||| |  ||||||||  ||  ||     || ||
Db        409 ACCTGCGATAGAGGCTGCCATCAAAAAGCTTGAAGATTCAGA 450


RESULT 12
US-10-767-701-29085
; Sequence 29085, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 29085
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 8548715
US-10-767-701-29085

  Query Match             11.6%;  Score 73.2;  DB 17;  Length 471;
  Best Local Similarity   58.1%;  Pred. No. 4.8e-12;
  Matches  129;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy        135 CGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGAC 194
              || | | |  ||  |  || |   ||   | |||||  || |||||||||| |  ||  |
Db         37 CGACCAAGGAGACTTTATCCTGTGCCGCCTGTCGAACAAGAGGAGGGTGACTCTGTCGGA 96

Qy        195 CTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCT 254
              |  || |||| |    | ||   |||||||||||| |||| ||||||  ||||||     |
Db         97 GTTCAAAGGCAGGTCACTGGTGTCCATCCGCGAGTTTTACGTGAAGGATGGCAAGGAGAT 156

Qy        255 CCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATAT 314
              |||  |  |||||||||     |  | ||| | ||||||  |   |  |  |   | |
Db        157 GCCCTCCGCCAAAGGTATTAGTATGACGATGTAGCAGTGGGAAGCATTTTGCAATGCTGT 216
```

```
Qy        315 CAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
              || |||||  | ||||| ||  ||  ||     || ||
Db        217 ACCTGCAATAGAGTATGCCATAAAAAAGTTTGAAGATTCAGA 258
```

RESULT 13
US-10-424-599-127795/c
; Sequence 127795, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 127795
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_86402C.1
US-10-424-599-127795

```
  Query Match             10.7%;  Score 67.4;  DB 16;  Length 497;
  Best Local Similarity   57.3%;  Pred. No. 3.2e-10;
  Matches  122;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        139 GACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGG 198
              || | |||| || || ||  |  | | ||  | || || |||||| || | | |   |
Db        474 GAAGGCGATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTGACGATTCAGGATTTC 415

Qy        199 AACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCC 258
              |  || ||      | |||  ||| || ||||  ||  | ||||| ||||||   || ||
Db        414 AGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGATGGCAAGGAGCTTCCT 355

Qy        259 GGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAA 318
                |||||| |||    |  ||   || ||||||   |   | | | | ||| |
Db        354 ACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAACCTTTAAGAAAAATGTGCCT 295

Qy        319 GCTATAGATGAGGCCATCAAGGAGAATGCGTGA 351
              || |||||  | ||||| ||| | |  | || |
Db        294 GCCATAGAAAAAGCCATTAAGAAAATGGAGTCA 262
```

RESULT 14
US-10-425-114-20666
; Sequence 20666, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20666
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-699-H3_FLI
US-10-425-114-20666

```
  Query Match             10.5%;  Score 66.2;  DB 16;  Length 794;
  Best Local Similarity   55.4%;  Pred. No. 9.5e-10;
  Matches  128;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        139 GACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGG 198
              || | |||| || || ||  |  | | ||  | || || |||||| || | | |   |
Db        373 GAAGGCGATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTGACGATTCAGGATTTC 432

Qy        199 AACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCC 258
              |  || ||      | |||  ||| || ||||  ||  | ||||| ||||||   || ||
Db        433 AGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGATGGCAAGGAACTTCCT 492

Qy        259 GGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAA 318
                |||||| |||    |  ||   || ||||||       | | | | ||| |
Db        493 ACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAGCCTTTAAGAAAAATGTGCCT 552

Qy        319 GCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTG 369
              || |||||  | ||||| ||| | |  | || |      | |||    |||
Db        553 GCCATAGAAAAAGCCATTAAGAAAATGGAGTCAAGTTGACACATGGCTTTG 603
```

RESULT 15
US-10-424-599-4356
; Sequence 4356, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 4356
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10393C.1
US-10-424-599-4356

```
  Query Match             10.5%;  Score 66.2;  DB 16;  Length 911;
  Best Local Similarity   55.4%;  Pred. No. 1e-09;
  Matches  128;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        139 GACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGG 198
              || | |||| || || ||  |  | | ||  | || || |||||| || | | |   |
Db        443 GAAGGCGATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTGACGATTCAGGATTTC 502

Qy        199 AACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCC 258
              |  || ||      | |||  ||| || ||||  ||  | ||||| ||||||   || ||
Db        503 AGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGATGGCAAGGAACTTCCT 562

Qy        259 GGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAA 318
                |||||| |||    |  ||   || ||||||       | | | | ||| |
Db        563 ACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAGCCTTTAAGAAAAATGTGCCT 622

Qy        319 GCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTG 369
              || |||||  | ||||| ||| | |  | || |      | |||    |||
Db        623 GCCATAGAAAAAGCCATTAAGAAAATGGAGTCAAGTTGACACATGGCTTTG 673
```

Search completed: January 11, 2005, 23:15:46
Job time : 486 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 11, 2005, 19:53:20 ; Search time 2756 Seconds
(without alignments)
8356.281 Million cell updates/sec

Title: US-10-629-953-3
Perfect score: 632
Sequence: 1 ctcctccaagactccacaag.....taaaaaaaaaaaaaaaaaa 632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	609.2	96.4	697	6	CB684473	CB684473 OSJNEf14A
c	2	609.2	96.4	784	6	CB684474	CB684474 OSJNEf14A
	3	608.8	96.3	693	1	AU082888	AU082888 AU082888
	4	501.4	79.3	539	1	AU092741	AU092741 AU092741
	5	453.2	71.7	530	7	D15350	D15350 RICC0499A R
	6	427.8	67.7	555	4	BI795760	BI795760 H030F04 E
	7	387.6	61.3	401	1	AU068340	AU068340 AU068340
	8	321.2	50.8	587	9	CL739643	CL739643 OR_BBa007
	9	277.6	43.9	643	8	AQ158085	AQ158085 nbxb0010L
	10	268	42.4	396	6	C27852	C27852 C27852 Rice
	11	262.2	41.5	329	2	BE607421	BE607421 OE202A06
	12	254.4	40.3	567	8	AQ795628	AQ795628 nbxb0056E
	13	241.6	38.2	631	6	CA151445	CA151445 SCJFRZ201
	14	240	38.0	443	6	C28075	C28075 C28075 Rice
	15	240	38.0	523	6	CA116904	CA116904 SCAGLR102
	16	240	38.0	649	6	CA189554	CA189554 SCCCLR1C0
	17	239.6	37.9	292	1	AU183581	AU183581 AU183581
	18	236.8	37.5	440	7	CF760736	CF760736 DSAF1_65_
	19	236.8	37.5	635	6	CA230818	CA230818 SCJFFL3C0
	20	234.2	37.1	855	6	CA129813	CA129813 SCUTLR201
	21	234	37.0	637	6	CA259070	CA259070 SCEQRT301
c	22	231.8	36.7	280	5	BX899341	BX899341 BX899341
	23	224	35.4	564	6	CD863080	CD863080 AZO1.105K
	24	222.4	35.2	606	6	CA189705	CA189705 SCCCLR1C0
	25	222.2	35.2	470	6	CD913769	CD913769 G550.119C
	26	221.6	35.1	403	2	BE424113	BE424113 WHE0076_G
	27	220.8	34.9	532	5	BQ838824	BQ838824 WHE3593_B
c	28	220.6	34.9	1072	7	CK212042	CK212042 FGAS02390
	29	220.4	34.9	460	6	CA710341	CA710341 wdk2c.pk0
	30	218.8	34.6	579	5	BU998906	BU998906 HI12I13r
c	31	218.8	34.6	618	6	CB860066	CB860066 HI12I13w
c	32	218.8	34.6	618	6	CB860077	CB860077 HI12J02w
	33	217.2	34.4	619	5	BU998918	BU998918 HI12J02r
	34	211.8	33.5	288	1	AU068341	AU068341 AU068341
	35	209.6	33.2	568	6	CB250145	CB250145 3529_1_18
c	36	209.6	33.2	596	6	CD568912	CD568912 3529_1_99
	37	206.8	32.7	586	6	CD568913	CD568913 3529_1_99
c	38	203.8	32.2	618	1	AI734589	AI734589 606033E08
c	39	203.6	32.2	603	1	AI691445	AI691445 606018A12
c	40	202.6	32.1	566	2	BE186543	BE186543 946008A01
	41	201.2	31.8	516	6	CF019517	CF019517 QBN1d08.x
c	42	201.2	31.8	588	7	CF625939	CF625939 zmrws05_0
c	43	201.2	31.8	588	7	CF627159	CF627159 zmrws05_0
	44	199	31.5	652	3	AY112536	AY112536 Zea mays
	45	197.6	31.3	553	9	CL715353	CL715353 OR_BBa004

ALIGNMENTS

RESULT 1
CB684473
LOCUS CB684473 697 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEf14A12.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA clone OSJNEf14A12 5', mRNA sequence.
ACCESSION CB684473
VERSION CB684473.1 GI:29688198
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 697)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 14 row: A column: 12
Seq primer: gta aaa cga cgg cca gtg.
FEATURES Location/Qualifiers
source 1. .697
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEf14A12"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEf"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Uninfected Control"
ORIGIN

```
  Query Match           96.4%;  Score 609.2;  DB 6;  Length 697;
  Best Local Similarity   99.5%;  Pred. No. 4.4e-155;
  Matches  611;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 CTCCTCCTAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 69

Qy         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 129

Qy        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 189

Qy        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 249

Qy        241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        250 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGATA 309

Qy        301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 369

Qy        361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 429

Qy        421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        430 TGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 489

Qy        481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 549

Qy        541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 609

Qy        601 GGTTTCTGCGTGTA 614
              ||||||||||||||
Db        610 GGTTTCTGCGTGTA 623


RESULT 2
CB684474/c
LOCUS       CB684474                784 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEf14A12.r OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf14A12 3', mRNA sequence.
ACCESSION   CB684474
VERSION     CB684474.1  GI:29688199
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 14  row: A  column: 12
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf14A12"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

  Query Match           96.4%;  Score 609.2;  DB 6;  Length 784;
  Best Local Similarity   99.5%;  Pred. No. 4.5e-155;
  Matches  611;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 626

Qy         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 566

Qy        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 506

Qy        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTTTACGAGAAG 446

Qy        241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        445 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGATA 386

Qy        301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 326

Qy        361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 266

Qy        421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        265 TGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 206

Qy        481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 146

Qy        541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 86

Qy        601 GGTTTCTGCGTGTA 614
              ||||||||||||||
Db         85 GGTTTCTGCGTGTA 72


RESULT 3
```

```
AU082888
LOCUS       AU082888                693 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AU082888 Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C53889, mRNA sequence.
ACCESSION   AU082888
VERSION     AU082888.1  GI:7212585
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from callus (2000)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C53889"
                     /clone_lib="Rice callus"
                     /note="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
                     NotI; cDNA prepared from rice callus mRNAs by using
                     oligo(dT) as a primer and ligating to the SalI-NotI site
                     of pBluescript II SK+ phagemid.  "
ORIGIN

  Query Match            96.3%;  Score 608.8;  DB 1;  Length 693;
  Best Local Similarity  99.3%;  Pred. No. 5.7e-155;
  Matches  610;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||||||||||||||||||||||||||||||| |||||||||||| |||||||||
Db         17 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTNTGAGAAGAATGTGNCGGAAGGGG 76

Qy         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 136

Qy        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 196

Qy        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAG 256

Qy        241 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATA 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        257 GACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGATA 316

Qy        301 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCC 376

Qy        361 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATAT 436

Qy        421 TGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 480
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        437 TGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGT 496

Qy        481 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 TTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGC 556

Qy        541 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 CTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTT 616

Qy        601 GGTTTCTGCGTGTA 614
              ||||||||||||||
Db        617 GGTTTCTGCGTGTA 630


RESULT 4
AU092741
LOCUS       AU092741                539 bp    mRNA    linear   EST 03-APR-2002
DEFINITION  AU092741 Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C53228, mRNA sequence.
ACCESSION   AU092741
VERSION     AU092741.1  GI:8527926
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from callus (2000)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C53228"
                     /clone_lib="Rice callus"
                     /note="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
                     NotI; cDNA prepared from rice callus mRNAs by using
                     oligo(dT) as a primer and ligating to the SalI-NotI site
                     of pBluescript II SK+ phagemid.  "
ORIGIN

  Query Match            79.3%;  Score 501.4;  DB 1;  Length 539;
  Best Local Similarity  96.7%;  Pred. No. 1.1e-125;
  Matches  523;  Conservative    0;  Mismatches   16;  Indels    2;  Gaps    1;

Qy         81 CGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGA 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGA 60

Qy        141 CGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAA 200
              |||||||||||||||||||||||||||||||||||||||| ||||   |  |  | |
Db         61 CGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGCCGGT--CGGACCTGGA 118

Qy        201 CGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGG 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 CGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGG 178

Qy        261 CCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGC 320
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        179 CCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGC 238
```

```
Qy        321 TATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTGTGATGCAAGTA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCCCATTCTCTTGTGATGCAAGTA 298

Qy        381 GACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATATTGCACTTTTGGTATGGTAAT 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATATTGCACTTTTGGTATGGTAAT 358

Qy        441 CTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGTTTGGTCATGCGCGGAAAATG 500
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 CTTCATTACCCTGTCTAGTAATATCAAGTAATGTTTTAGTTTGGTCATGCGCGGAAAATG 418

Qy        501 TGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGCCTAATGTGCTCTCCTCATTT 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTATGTGCCTAATGTGCTCTCCTCATTT 478

Qy        561 TGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTTGGTTTCTGCGTGTAAAAAAA 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTTGGTTGGTTTCAAAAAAAAAAAAAA 538

Qy        621 A 621
              |
Db        539 A 539
```

RESULT 5
D15350
LOCUS D15350 530 bp mRNA linear EST 24-FEB-2003
DEFINITION RICC0499A Rice callus Oryza sativa (japonica cultivar-group) cDNA clone C0499, mRNA sequence.
ACCESSION D15350 C98059
VERSION D15350.2 GI:15072274
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 530)
AUTHORS Sasaki,T., Song,J., Koga-Ban,Y., Matsui,E., Fang,F., Higo,H., Nagasaki,H., Hori,M., Miya,M. and Murayama-Kayano,E.
TITLE Toward cataloguing all rice genes: large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library
JOURNAL Plant J. 6 (4), 615-624 (1994)
MEDLINE 95078950
PUBMED 7987417
COMMENT On May 17, 1993 this sequence version replaced gi:286543 gi:3760805.
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT ='RGP'
Seq primer: oligo(dT).
FEATURES Location/Qualifiers
source 1. .530
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="C0499"
/clone_lib="Rice callus"
/note="Vector: pBluescript II SK+; Site_1: SalI; Site_2: NotI; cDNA prepared from rice callus mRNAs by using oligo(dT) as a primer and ligating to the SalI-NotI site of pBluescript II SK+ phagemid. "
ORIGIN

Query Match 71.7%; Score 453.2; DB 7; Length 530;
Best Local Similarity 97.5%; Pred. No. 1.6e-112;
Matches 499; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

```
Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              |||||||||||||||||||||||| |||||||||||||||||||| ||||||||||
Db         21 CTCCTCCAAGACTCCACAAGTCCAGANGGAGAAAGTGTGAGAAGAATGNGGCGGAAGGGG 80

Qy         61 AACAAGCGGTTCGG-CGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGA 119
              ||||||||||||| |||||||||||| ||| ||||||||||||||||||||||||||||
Db         81 AACAAGCGGTTCGGACGGCGGCGGCGANCCGACGGCCAAGCGCCGTGCCGCCGGGGACGA 140

Qy        120 CGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAG 179
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        141 CGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTNGCCCAGATATCGAAGAACAGGAG 200

Qy        180 GGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAA 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAA 260

Qy        240 GGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAAGAT 299
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        261 GGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAAGAT 320

Qy        300 ACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGAGCC 359
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        321 ACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAG-ATGCGTGATCGGAGCC 379

Qy        360 CATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATA 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGAATA 439

Qy        420 TTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTTTAG 479
              ||||||||||||||| ||||||||| | ||||||||||||||||||||||||||||||
Db        440 TTGCACTTTTGGTANGGTAATCTTCA-TACCCTGTCTAGTAATATCAAGTAATGTTTTAG 498

Qy        480 TTTGGTCATGCGCGG-AAAATGTGGCCTTGTC 510
              |||||||| |||||| ||||||||||||||||
Db        499 TTTGGTCANGCGCGGAAAAATGTGGCCTTGTC 530
```

RESULT 6
BI795760
LOCUS BI795760 555 bp mRNA linear EST 02-OCT-2001
DEFINITION H030F04 Endosperm library from Oryza sativa (10 days after anthesis) Oryza sativa cDNA clone H030F04, mRNA sequence.
ACCESSION BI795760
VERSION BI795760.1 GI:15847484
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 555)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Debao Li
Bioinfomatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES Location/Qualifiers
source 1. .555
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="H030F04"
/tissue_type="Endosperm"
/dev_stage="10 days after anthesis"

/clone_lib="Endosperm library from Oryza sativa (10 days after anthesis)"
/note="Vector: pSport2"
ORIGIN

Query Match 67.7%; Score 427.8; DB 4; Length 555;
Best Local Similarity 94.6%; Pred. No. 1.3e-105;
Matches 521; Conservative 0; Mismatches 17; Indels 13; Gaps 7;

```
Qy         53 GGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCG 112
              | ||||||||||||||| |    || ||    ||||||||| ||||||| | |    | |||
Db          1 GAAAGGGGAACAAGCGTTCGCGCCGC----GCGAGCCGGCCGCCAAGCCCTG---CCCCG 53

Qy        113 GGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGA 172
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         54 GGGACGACGGGCCCTCCGAGAGC-CCGACGACGATATCGTCGTCGCCCAGATATCGAAGA 112

Qy        173 ACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTC-GTCGTCGACATCCGCGAGTTC 231
              |||||||||||    | |||||||||||||||||||| |||||||||||||||||||||
Db        113 ACAGGAGGGTG--CGGTCGGACCTGGAACGGCAAGGTCTGTCGTCGACATCCGCGAGTTC 170

Qy        232 TACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATAC-AGCTCCCAATGGATCA 290
              |||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||
Db        171 TACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAAGCTCCCGATGGATCA 230

Qy        291 GTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTG 350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTG 290

Qy        351 ATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTC 410
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        291 ATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTTCGTCTGTCTTTTATGACTC 350

Qy        411 CGAGGAATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTA 470
              |||||||||||||||||||||||||||||| | ||||||||||||||||||||||||||
Db        351 CGAGGAATATTGCACTTTTGGTATGGTAATCTTCACTACCCTGTCTAGTAATATCAAGTA 410

Qy        471 ATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGC 530
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        411 ATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCC-AAGAAGC 469

Qy        531 AGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTA 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 AGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTA 529

Qy        591 ATGGTTGGTTG 601
              ||||||||||||
Db        530 ATGGTTGGTTG 540
```

RESULT 7
AU068340
LOCUS AU068340 401 bp mRNA linear EST 02-APR-2002
DEFINITION AU068340 Rice callus Oryza sativa (japonica cultivar-group) cDNA clone C20066_11Z, mRNA sequence.
ACCESSION AU068340
VERSION AU068340.1 GI:5005023
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 401)
AUTHORS Yamamoto,K. and Sasaki,T.
TITLE Rice cDNA from callus (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT ='RGP'.
FEATURES Location/Qualifiers
source 1. .401
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="C20066_11Z"
/clone_lib="Rice callus"
/note="Vector: pBluescript II SK+; Site_1: SalI; Site_2: NotI; cDNA prepared from rice callus mRNAs by using oligo(dT) as a primer and ligating to the SalI-NotI site of pBluescript II SK+ phagemid. "
ORIGIN

Query Match 61.3%; Score 387.6; DB 1; Length 401;
Best Local Similarity 98.5%; Pred. No. 1.2e-94;
Matches 390; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy        237 GAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATGGATCAGTGGAA 296
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 GAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATGGATCAGTGGAA 60

Qy        297 GATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCGTGATCGGA 356
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         61 GATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATNCGTGATCGGA 120

Qy        357 GCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGACTCCGAGGA 416
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        121 GCCCATTCTCTTGTGATGCAAGTAGNCTAAGCCTACGTCTGTCTTTTATGACTCCGAGGA 180

Qy        417 ATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAGTAATGTTT 476
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 ATATTGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAGTAATGTTT 240

Qy        477 TAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTAT 536
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAAGCAGCTAT 300

Qy        537 GTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTT 596
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGTTAATGGTT 360

Qy        597 GGTTGGTTTCTGCGTGTAAAAAAAAAAAAAAAAAAA 632
              |||||||||||||||||  ||||||||||||||||
Db        361 GGTTGGTTTCTGCGTGTACTAAAAAAAAAAAAAAAA 396
```

RESULT 8
CL739643
LOCUS CL739643 587 bp DNA linear GSS 27-JUL-2004
DEFINITION OR_BBa0074L07.f OR_BBa Oryza rufipogon genomic clone OR_BBa0074L07 5', genomic survey sequence.
ACCESSION CL739643
VERSION CL739643.1 GI:50674979
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 587)
AUTHORS Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259

```
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 161   Std Error: 0.00
            Plate: 0074  row: L  column: 07
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .587
                      /organism="Oryza rufipogon"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4529"
                      /clone="OR_BBa0074L07"
                      /tissue_type="young leaves"
                      /lab_host="DH10B-T1 phage resistant"
                      /clone_lib="OR_BBa"
                      /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             50.8%;  Score 321.2;  DB 9;  Length 587;
  Best Local Similarity   99.1%;  Pred. No. 1.7e-76;
  Matches  323;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        289 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 295

Qy        349 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 355

Qy        409 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAG 468
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        356 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAG 415

Qy        469 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 475

Qy        529 GCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTGT 588
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        476 GCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGGTGGGTAAGGAGGTCCATCTGT 535

Qy        589 TAATGGTTGGTTGGTTTCTGCGTGTA 614
              |||||||||| ||||||||||||||||
Db        536 TAATGGTTGGGTGGTTTCTGCGTGTA 561


RESULT 9
AQ158085
LOCUS       AQ158085                643 bp    DNA     linear   GSS 12-SEP-1998
DEFINITION  nbxb0010L09r CUGI Rice BAC Library Oryza sativa (japonica
            cultivar-group) genomic clone nbxb0010L09r, genomic survey
            sequence.
ACCESSION   AQ158085
VERSION     AQ158085.1  GI:3555110
KEYWORDS    GSS.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 25
            High quality sequence stop: 517.
FEATURES              Location/Qualifiers
     source           1. .643
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="genomic DNA"
                      /cultivar="Nipponbare"
                      /db_xref="taxon:39947"
                      /clone="nbxb0010L09r"
                      /tissue_type="Leaf"
                      /lab_host="E. coli DH10B"
                      /clone_lib="CUGI Rice BAC Library"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII; Rice is one of two most popular grains in the
                      world. Half of the world  population especially those
                      inhabiting highly populated areas of the  humid tropics
                      and subtropics, rely on rice as their primary source of
                      carbohydrate. Monocotyledonous rice is a diploid plant
                      (2n=24) with a  haploid genome equivalent of 431 Mbp
                      (Arumuganathan and Earle, 1991).  The relatively small
                      genome of rice, three times larger than that of
                      Arabidopsis, makes it suitable for genomic studies. In
                      order to  facilitate positional cloning, physical mapping
                      and genome sequencing  of rice, we have constructed a BAC
                      library from Oryza sativa, Nipponbare  variety. The
                      library contains 36,864 clones with an average insert size
                      of 128.5 Kb providing 10.9 haploid genome equivalents. The
                      deep  coverage allows the isolation a particular sequence
                      with a probability  of 99.9 %. Two high density filters,
                      each containing 18,432 clones  (doubly spotted), represent
                      the whole library for colony screening."
ORIGIN

  Query Match             43.9%;  Score 277.6;  DB 8;  Length 643;
  Best Local Similarity   98.3%;  Pred. No. 1.3e-64;
  Matches  291;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        289 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 407

Qy        349 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 467

Qy        409 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAG 468
              ||||||||||||||| ||||| |||||||||||| ||||||||||||||||||||||||
Db        468 TCCGAGGAATATTGCTCTTTT-GTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAG 526

Qy        469 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 586

Qy        529 GCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCAT 584
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||
Db        587 GCAGCTATGTGCCTAATGTGCTCTCCTCATTATGTACGTTGTGTAAGGAGGTCCAT 642


RESULT 10
C27852
LOCUS       C27852                  396 bp    mRNA    linear   EST 04-APR-2002
DEFINITION  C27852 Rice callus cDNA Oryza sativa (japonica cultivar-group) cDNA
            clone C53228_1A, mRNA sequence.
ACCESSION   C27852
VERSION     C27852.1  GI:2311697
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

```
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Yamamoto,K. and Sasaki,T.
  TITLE     Rice cDNA from callus 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C53228_1A"
                     /tissue_type="callus"
                     /dev_stage="callus"
                     /clone_lib="Rice callus cDNA"
ORIGIN

  Query Match             42.4%;  Score 268;  DB 6;  Length 396;
  Best Local Similarity   88.1%;  Pred. No. 5.1e-62;
  Matches  342;  Conservative    0;  Mismatches   37;  Indels    9;  Gaps    5;

Qy          6 CCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGGAACAA 65
              ||||||| ||||||||||||||||||||||||||| |||||||| | || ||||||||||||
Db          1 CCAAGACNCCACAAGTCCAGAGGGAGAAAGTGTNAGAAGAATNTAGCNGAAGGGGAACAA 60

Qy         66 GCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGACGGGCC 125
              |||||||  | ||  | || |||  |     | || |||  || | | | ||||||||||
Db         61 GCGGTTCNACAGCANCAGCAAGCGCGGCCAGATGCACCGNACCACGGNGNACGACGGGCC 120

Qy        126 CTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGGGTGGC 185
              ||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db        121 CTCCGAGAGCGCCGACGACGATATCGTC--TNCCCAGATATCGAAGAACAGGAGGGTGGC 178

Qy        186 GGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGG 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 GGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGAGAAGGACGG 238

Qy        246 CAAGACCCTCCCCGGCCGCAAAGGTA-TACAGCTCCCAATGGATCAGTGGAAGATACTGA 304
              |||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        239 CAAGACCCTCCCCGGCCGCAAAGGTATTACAGCTCCCGATGGATCAGTGGAAGATACTTN 298

Qy        305 --GGGACAATATCAAAGCTATAGATGAGGCCATCAAGGA-GAATGCGTGATCGGAGCCCA 361
                |||||||||||||||||||||||||||||||||||||| |||||||| ||||||||||
Db        299 AGGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGGAATGCGTNATCGGAGCCCA 358

Qy        362 ---TTCTCTTGTGATGCAAGTAGACTAA 386
                 |||| |  |||||||||||||| |
Db        359 TTTTTCTTTGTTGATGCAAGTAGACTTA 386


RESULT 11
BE607421
LOCUS       BE607421                329 bp    mRNA    linear   EST 23-AUG-2000
DEFINITION  OE202A06 OE Oryza sativa (indica cultivar-group) cDNA clone Rice
            Microarray ID699 similar to unknown, mRNA sequence.
ACCESSION   BE607421
VERSION     BE607421.1  GI:9886329
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M.,
            Ferrea,H., Kawasaki,S., McCollough,A., Michalowski,C.B.,
            Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.
  TITLE     Functional Genomics of Plant Stress Tolerance
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Michalowski,C.B.
            University of Arizona
            Bio Sciences West room 513, Tucson, AZ 85721, USA
            Tel: 520-621-7982
            Fax: 520-621-1697
            Email: cbm@u.arizona.edu
            No hits found using blastx vs. genbank:(nr), cutoff:(e=0.1)
            Unknown, putative, similar, and hypothetical are not characterized
            with biological features.
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /strain="Pokkali"
                     /db_xref="taxon:39946"
                     /clone="Rice Microarray ID699"
                     /tissue_type="roots"
                     /dev_stage="1 week"
                     /clone_lib="OE"
                     /note="2-3 d 150mM NaCl"
ORIGIN

  Query Match             41.5%;  Score 262.2;  DB 2;  Length 329;
  Best Local Similarity   98.9%;  Pred. No. 1.9e-60;
  Matches  264;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        348 GTGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGA 407
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          9 GTGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTTCGTCTGTCTTTTATGA 68

Qy        408 CTCCGAGGAATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAA 467
              |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||||
Db         69 CTCCGAGGAATATTGCACTTTTGGTATGGTAATCTTCACTACCCTGTCTAGTAATATCAA 128

Qy        468 GTAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGA 527
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GTAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGA 188

Qy        528 AGCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTG 587
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 AGCAGCTATGTGCCTAATGTGCTCTCCTCATTTTGTACGTTGGGTAAGGAGGTCCATCTG 248

Qy        588 TTAATGGTTGGTTGGTTTCTGCGTGTA 614
              ||||||||||||||||||||||||||||
Db        249 TTAATGGTTGGTTGGTTTCTGCGTGTA 275


RESULT 12
AQ795628
LOCUS       AQ795628                567 bp    DNA     linear   GSS 04-AUG-1999
DEFINITION  nbxb0056E24r CUGI Rice BAC Library Oryza sativa (japonica
            cultivar-group) genomic clone nbxb0056E24r, genomic survey
            sequence.
ACCESSION   AQ795628
VERSION     AQ795628.1  GI:5704013
KEYWORDS    GSS.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 567)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
```

```
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence stop: 494.
FEATURES              Location/Qualifiers
     source           1. .567
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="genomic DNA"
                      /cultivar="Nipponbare"
                      /db_xref="taxon:39947"
                      /clone="nbxb0056E24r"
                      /tissue_type="Leaf"
                      /lab_host="E. coli DH10B"
                      /clone_lib="CUGI Rice BAC Library"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII; Rice is one of two most popular grains in the
                      world. Half of the world  population especially those
                      inhabiting highly populated areas of the  humid tropics
                      and subtropics, rely on rice as their primary source of
                      carbohydrate. Monocotyledonous rice is a diploid plant
                      (2n=24) with a  haploid genome equivalent of 431 Mbp
                      (Arumuganathan and Earle, 1991).  The relatively small
                      genome of rice, three times larger than that of
                      Arabidopsis, makes it suitable for genomic studies. In
                      order to  facilitate positional cloning, physical mapping
                      and genome sequencing  of rice, we have constructed a BAC
                      library from Oryza sativa, Nipponbare  variety. The
                      library contains 36,864 clones with an average insert size
                      of 128.5 Kb providing 10.9 haploid genome equivalents. The
                      deep  coverage allows the isolation a particular sequence
                      with a probability  of 99.9 %. Two high density filters,
                      each containing 18,432 clones  (doubly spotted), represent
                      the whole library for colony screening."
ORIGIN

  Query Match            40.3%;  Score 254.4;  DB 8;  Length 567;
  Best Local Similarity  99.6%;  Pred. No. 2.8e-58;
  Matches  255;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        289 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAATGCG 370

Qy        349 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TGATCGGAGCCCATTCTCTTGTGATGCAAGTAGACTAAGCCTACGTCTGTCTTTTATGAC 430

Qy        409 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTTATTACCCTGTCTAGTAATATCAAG 468
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        431 TCCGAGGAATATTGCACTTTTGGTATGGTAATCTTCATTACCCTGTCTAGTAATATCAAG 490

Qy        469 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 TAATGTTTTAGTTTGGTCATGCGCGGAAAATGTGGCCTTGTCTGGTATTGTGCCAAAGAA 550

Qy        529 GCAGCTATGTGCCTAA 544
              ||||||||||||||||
Db        551 GCAGCTATGTGCCTAA 566


RESULT 13
CA151445
LOCUS       CA151445                631 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCJFRZ2010C02.g RZ2 Saccharum officinarum cDNA clone SCJFRZ2010C02
            5', mRNA sequence.
ACCESSION   CA151445
VERSION     CA151445.1  GI:35055505
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 010  row: C  column: 02
            Seq primer: T7 Promoter Primer.
FEATURES              Location/Qualifiers
     source           1. .631
                      /organism="Saccharum officinarum"
                      /mol_type="mRNA"
                      /db_xref="taxon:4547"
                      /clone="SCJFRZ2010C02"
                      /lab_host="DH10B"
                      /clone_lib="RZ2"
                      /note="Organ: Shoot-root transition zone from young plants
                      (small insert library); Vector: pSport1; Site_1: SalI;
                      Site_2: NotI; An unidirectional cDNA library generated
                      from [Shoot-root transition zone  from young plants (small
                      insert library)]. cDNA was prepared from polyA+  mRNA
                      using SuperScript Plasmid System Kit (Invitrogen). The
                      double-strand  cDNAs were fractionated in a sepharose
                      CL-2B 40cm-columns and fragments  sizing between 0.8 and
                      1.5 Kb were directionally cloned into the vector.  Details
                      of each source of RNA and library construction can be
                      obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match            38.2%;  Score 241.6;  DB 6;  Length 631;
  Best Local Similarity  84.7%;  Pred. No. 8.9e-55;
  Matches  271;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy         40 AGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAG 99
              || |  |||||| ||||||| || ||||| ||||||||||||||||||||||||||||||
Db          3 AGGAAGATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAG 62

Qy        100 CGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCC 159
              ||||  |||||||||||| ||||| |||||||||    |||||| |||| ||  ||||| |||
Db         63 CGCCAGGCCGCCGGGGATGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCC 122

Qy        160 CAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGAC 219
              |||||||||||||||| || |||| | ||  ||| |||||| |||| ||||  |||||||
Db        123 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAGGGCAGGGTCTACGTCGAC 182

Qy        220 ATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTC 279
              ||||||||||||||  |||||||||||||| ||||||  ||||||||||||   |||
Db        183 CTCCGCGAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 242

Qy        280 CCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAG 339
              |    | ||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        243 CAGTTAGATCAGTGGAAGATACTGAGGGACAATATCAAAGCCATAGATGAGGCCATCAAG 302

Qy        340 GAGAATGCGTGATCGGAGCC 359
              |||||  | |||| ||||||
Db        303 GAGAACACATGATGGGAGCC 322


RESULT 14
C28075
LOCUS       C28075                  443 bp    mRNA    linear   EST 04-APR-2002
DEFINITION  C28075 Rice callus cDNA Oryza sativa (japonica cultivar-group) cDNA
```

```
            clone C53889_2A, mRNA sequence.
ACCESSION   C28075
VERSION     C28075.1  GI:2311920
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Yamamoto,K. and Sasaki,T.
  TITLE     Rice cDNA from callus 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C53889_2A"
                     /tissue_type="callus"
                     /dev_stage="callus"
                     /clone_lib="Rice callus cDNA"
ORIGIN

  Query Match             38.0%;  Score 240;  DB 6;  Length 443;
  Best Local Similarity   92.7%;  Pred. No. 2.3e-54;
  Matches  303;  Conservative    0;  Mismatches   17;  Indels    7;  Gaps    5;

Qy          1 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 CTCCTCCAAGACTCCACAAGTCCAGAGGGAGAAAGTGTGAGAAGAATGTGGCGGAAGGGG 77

Qy         61 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 AACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGTGCCGCCGGGGACGAC 137

Qy        121 GGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATATCGAAGAACAGGAGG 180
              |||||||||||||| ||||||||||||||  |   ||||||||||||||||||||||.
Db        138 GGGCCCTCCGAGAGCNCCGACGACGATATCGT-NTNCNCCAGATATCGAAGAACAGGAGG 196

Qy        181 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAG-TTCTACGAGAA 239
              |||||||||||||||||||||||||||||||||||||||||||||| || || ||||||||
Db        197 GTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCNAGTTTNTACGAGAA 256

Qy        240 GGACGGCAAGACCCTCCCCGGCCGC-AAAGGTATACAGC-TCCCAATGGATCAGTGG--- 294
              |||||||||||||||||||||||||| |||||||||||||| |||| || |||||||  |
Db        257 GGACGGCAAGACCCTCCCCGGCCGCAAAAGGTATACAGCTTCCCGATNGATCAGTTGGAA 316

Qy        295 AAGATACTGAGGGACAATATCAAAGCT 321
              | |   |  ||||||||||||||||||
Db        317 GATACTTTAGGGGACAATATCAAAGCT 343


RESULT 15
CA116904
LOCUS       CA116904                523 bp    mRNA    linear   EST 23-SEP-2003
DEFINITION  SCAGLR1021A08.g LR1 Saccharum officinarum cDNA clone SCAGLR1021A08
            5', mRNA sequence.
ACCESSION   CA116904
VERSION     CA116904.1  GI:34970211
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 523)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 021  row: A  column: 08
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCAGLR1021A08"
                     /lab_host="DH10B"
                     /clone_lib="LR1"
                     /note="Organ: Leaf roll from field grown adult plants
                     (large insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Leaf roll from field grown  adult plants (large
                     insert library)]. cDNA was prepared from polyA+ mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             38.0%;  Score 240;  DB 6;  Length 523;
  Best Local Similarity   84.4%;  Pred. No. 2.3e-54;
  Matches  270;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy         40 AGAAGAATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAG 99
              ||  |  |||||| |||||| || ||||| ||||||||||||||||||||||||||||||
Db          6 AGGAAGATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAG 65

Qy        100 CGCCGTGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCC 159
              ||||  |||||||||| ||||| ||||||||    |||||| |||| ||  ||||| |||
Db         66 CGCCAGGCCGCCGGGGATGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCC 125

Qy        160 CAGATATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGAC 219
              |||||||||||||||  || |||| | ||  ||| ||||||  |||| ||||  |||||||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTACGTCGAC 185

Qy        220 ATCCGCGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTC 279
              ||||||||||||||  ||||||||||||| ||||||  ||||||||||||    |||
Db        186 CTCCGCGAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 245

Qy        280 CCAATGGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAG 339
              |   | |||||||||||||| ||||||||||||||||||| ||||||||||||||||||
Db        246 CAGTTAGATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGATGAGGCCATCAAG 305

Qy        340 GAGAATGCGTGATCGGAGCC 359
              |||||  | |||| ||||||
Db        306 GAGAACACATGATGGGAGCC 325


Search completed: January 11, 2005, 22:08:08
Job time : 2763 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:49:11 ; Search time 152 Seconds
(without alignments)
238.366 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	530	100.0	101	3	AAY44879	Aay44879 Rice tran
2	398	75.1	103	3	AAY44878	Aay44878 Corn tran
3	389.5	73.5	96	3	AAY44881	Aay44881 Wheat tra
4	269	50.8	105	3	AAY44880	Aay44880 Soybean t
5	247.5	46.7	107	3	AAG18457	Aag18457 Arabidops
6	247.5	46.7	142	3	AAG18455	Aag18455 Arabidops
7	206.5	39.0	111	3	AAY44883	Aay44883 Vernonia
8	194.5	36.7	85	3	AAG04501	Aag04501 Arabidops
9	194.5	36.7	120	3	AAG04499	Aag04499 Arabidops
10	185	34.9	165	6	ABR57409	Abr57409 Brassica
11	184.5	34.8	141	3	AAY44886	Aay44886 Soybean t
12	183	34.5	144	3	AAG50454	Aag50454 Arabidops
13	183	34.5	144	3	AAG09966	Aag09966 Arabidops
14	183	34.5	165	3	AAG50453	Aag50453 Arabidops
15	183	34.5	165	3	AAG09965	Aag09965 Arabidops
16	183	34.5	165	5	ABB92888	Abb92888 Herbicida
17	170.5	32.2	99	3	AAY44882	Aay44882 Marigold
18	165	31.1	140	3	AAY44885	Aay44885 Corn tran
19	165	31.1	183	3	AAY44884	Aay44884 Corn tran
20	160.5	30.3	127	4	ABB50267	Abb50267 Polymeras
21	160.5	30.3	127	6	ABR47556	Abr47556 Breast ca
22	160.5	30.3	127	7	ADF76608	Adf76608 Novel hum
23	160.5	30.3	127	8	ADL91494	Adl91494 Human imm
24	160.5	30.3	127	8	ADN04676	Adn04676 Antipsori
25	160.5	30.3	132	3	AAB54327	Aab54327 Human pan
26	154.5	29.2	150	4	ABG27269	Abg27269 Novel hum
27	145.5	27.5	110	4	ABB63838	Abb63838 Drosophil
28	144	27.2	127	8	ADN99783	Adn99783 Novel hum
29	131.5	24.8	128	5	ABP64800	Abp64800 Human pro
30	116	21.9	113	4	ABG23001	Abg23001 Novel hum
31	116	21.9	150	4	ABG26860	Abg26860 Novel hum
32	112.5	21.2	54	8	ADP29939	Adp29939 Human sec
33	112.5	21.2	54	8	ADP30051	Adp30051 Human sec
34	110	20.8	104	5	ABP33375	Abp33375 Human RNA
35	106	20.0	193	4	ABG23003	Abg23003 Novel hum
36	105.5	19.9	208	4	ABG26861	Abg26861 Novel hum
37	100	18.9	234	3	AAG06282	Aag06282 Arabidops
38	100	18.9	245	3	AAG06281	Aag06281 Arabidops
39	98	18.5	72	3	AAG14414	Aag14414 Arabidops
40	98	18.5	95	3	AAG14413	Aag14413 Arabidops
41	82.5	15.6	11877	8	ADL91934	Adl91934 Streptomy
42	82.5	15.6	12199	3	AAY77180	Aay77180 S. venezu
43	79.5	15.0	161	3	AAG06283	Aag06283 Arabidops
44	79.5	15.0	336	3	AAG11256	Aag11256 Arabidops
45	79.5	15.0	339	4	ABG26859	Abg26859 Novel hum

ALIGNMENTS

RESULT 1
AAY44879
ID AAY44879 standard; protein; 101 AA.
XX
AC AAY44879;
XX
DT 18-MAY-2000 (first entry)
XX
DE Rice transcription coactivator PC4(P15) type 1.
XX
KW Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW rice; signal mediator; activator; Gal4/VP16; transgenic plant;
KW general transcription factor; GTF; transcription initiation complex;
KW immunological screening; detection; marker; clone rr1.pk0003.a12.
XX
OS Oryza sativa.
XX
PN WO200005377-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016479.
XX
PR 22-JUL-1998; 98US-0093687P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Caimi PG, Odell JT, Sakai H, Zhu Q;
XX
DR WPI; 2000-182701/16.
DR N-PSDB; AAZ50412.
XX
PT Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT polynucleotides and polypeptides.
XX
PS Claim 12; Page 33; 43pp; English.
XX
CC The present amino acid sequence is the rice Positive Cofactor 4 (PC4)
CC transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC from clone rr1.pk0003.a12, obtained from rr1 cDNA library prepared using
CC rice root two week old developing seedlings. This sequence has 45%
CC homology to PC4(P15) type 1 from Arabidopsis thaliana (gi 2997684). PC4
CC molecules functions as a signal mediator between activators like
CC Gal4/VP16 and general transcription factors (GTFs) in a transcription
CC initiation complex. The expression levels of PC4 can be manipulated and
CC the functional properties of specific transcriptional activators can be
CC modulated. The PC4 DNA can be used to create transgenic plants with

```
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 101 AA;

  Query Match             100.0%;  Score 530;  DB 3;  Length 101;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-57;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60

Qy         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
              |||||||||||||||||||||||||||||||||||||||||
Db         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101


RESULT 2
AAY44878
ID   AAY44878 standard; protein; 103 AA.
XX
AC   AAY44878;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Corn transcription coactivator PC4(P15) type 1.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   corn; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone cca.pk0020.d2.
XX
OS   Zea mays.
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   N-PSDB; AAZ50411.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 6; Page 32-33; 43pp; English.
XX
CC   The present amino acid sequence is the corn Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from the corn callus type II tissue (undifferentiated) from the cDNA
CC   clone cca.pk0020.d2. This sequence has 38% homology to PC4(P15) type 1
CC   from Arabidopsis thaliana (gi 2997684). PC4 molecules functions as a
CC   signal mediator between activators like Gal4/VP16 and general
CC   transcription factors (GTFs) in a transcription initiation complex. The
CC   expression levels of PC4 can be manipulated and the functional properties
CC   of specific transcriptional activators can be modulated. The PC4 DNA can
CC   be used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 103 AA;

  Query Match             75.1%;  Score 398;  DB 3;  Length 103;
  Best Local Similarity   74.5%;  Pred. No. 5.8e-41;
  Matches   76;  Conservative   13;  Mismatches   11;  Indels    2;  Gaps    1;

Qy          1 MWRKGNKRFGGG--GEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVD 58
              || || ||||||  |||||||:|| ||||||||:|  |||:||||::|:||:| |:| ||
Db          1 MWGKGKKRFGGGGSGEPAAKRQAARDDGPSESAEDGTVVAEISKNKKVSVRSWKGRVFVD 60

Qy         59 IREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKEN 100
              :|||| ||||||| |||| | :||||||:||||||:|||:||
Db         61 LREFYFKDGKTLPTRKGISLQLDQWKILKDNIKAINEAIEEN 102


RESULT 3
AAY44881
ID   AAY44881 standard; protein; 96 AA.
XX
AC   AAY44881;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Wheat transcription coactivator PC4(P15) type 1.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   wheat; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone wdk2c.pk015.g20.
XX
OS   Triticum aestivum.
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   N-PSDB; AAZ50414.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 6; Page 35; 43pp; English.
XX
CC   The present amino acid sequence is the wheat Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone wdk2c.pk015.g20, obtained from wdk2c cDNA library, prepared
CC   from wheat developing kernels, 7 days after anthesis. This sequence has
CC   45% homology to PC4(P15) type 1 from Arabidopsis thaliana (gi 2997684).
CC   PC4 molecules functions as a signal mediator between activators like
CC   Gal4/VP16 and general transcription factors (GTFs) in a transcription
CC   initiation complex. The expression levels of PC4 can be manipulated and
CC   the functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 96 AA;
```

```
  Query Match           73.5%;  Score 389.5;  DB 3;  Length 96;
  Best Local Similarity  80.0%;  Pred. No. 5.9e-40;
  Matches   76;  Conservative    7;  Mismatches    9;  Indels    3;  Gaps    1;

Qy          9 FGGG---GEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEK 65
              ||||   |:| |||:||| |||||  || |||||||||:||||| |||||:||:||||||
Db          1 FGGGGGRGQPPAKRQAAGKDGPSEETDDGIVVAQISKNKRVAVRNWNGKVMVDMREFYEK 60

Qy         66 DGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKEN 100
              |||:|| |||| | ||||||||||:|||||||||
Db         61 DGKSLPTRKGISLSMDQWKILRDNIEAIDEAIKEN 95


RESULT 4
AAY44880
ID   AAY44880 standard; protein; 105 AA.
XX
AC   AAY44880;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Soybean transcription coactivator PC4(P15) type 1.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   soybean; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone sfl1.pk0008.a4.
XX
OS   Glycine max.
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   N-PSDB; AAZ50413.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 6; Page 34; 43pp; English.
XX
CC   The present amino acid sequence is the soybean Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone sfl1.pk0008.a4, obtained from sfl1 cDNA library, prepared from
CC   soybean immature flowers. This sequence has 59% homology to PC4(P15) type
CC   1 from Arabidopsis thaliana (gi 2997684). PC4 molecules functions as a
CC   signal mediator between activators like Gal4/VP16 and general
CC   transcription factors (GTFs) in a transcription initiation complex. The
CC   expression levels of PC4 can be manipulated and the functional properties
CC   of specific transcriptional activators can be modulated. The PC4 DNA can
CC   be used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 105 AA;

  Query Match           50.8%;  Score 269;  DB 3;  Length 105;
  Best Local Similarity  54.4%;  Pred. No. 5.4e-25;
  Matches   56;  Conservative   13;  Mismatches   14;  Indels   20;  Gaps    3;

Qy         17 AKRRAAGDDGPSES---------------ADDD---IVVAQISKNRRVAVRTWNGKVVVD 58
              ||||   ||| |::               :|||   | | :|||||||||| | | ::||
Db          5 AKRR--DDDGASDADSEGHAPPKKSLKKDSDDDPDSITVCEISKNRRVAVRNWKGSIMVD 62

Qy         59 IREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
              ||||| |||| ||||||| | |||| :||:::: ||:|| ||:
Db         63 IREFYVKDGKQLPGRKGISLTMDQWNVLRNHVEEIDKAINENS 105


RESULT 5
AAG18457
ID   AAG18457 standard; protein; 107 AA.
XX
AC   AAG18457;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 19873.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
```

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.7%; Score 247.5; DB 3; Length 107;
Best Local Similarity 56.5%; Pred. No. 2.6e-22;
Matches 48; Conservative 12; Mismatches 22; Indels 3; Gaps 1;

```
Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
              || |      ||       |||||  |||||||:|| |||| : :|||||| ||||||||:|
Db         24 PAKKVAKPADD---SDQSDDIVVCNISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKK 80
```

```
Qy          75 GIQLPMDQWKILRDNIKAIDEAIKE 99
               || | :|||  ||:: : |::|: :
Db          81 GISLSVDQWNTLRNHAEDIEKALSD 105
```

RESULT 6
AAG18455
ID AAG18455 standard; protein; 142 AA.
XX
AC AAG18455;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 19871.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

```
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             46.7%;  Score 247.5;  DB 3;  Length 142;
  Best Local Similarity   56.5%;  Pred. No. 3.8e-22;
  Matches   48;  Conservative   12;  Mismatches   22;  Indels    3;  Gaps    1;

Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
              || |     ||       |||||  |||||||:|| |||| : :|||||| ||||||||:|
Db         59 PAKKVAKPADD---SDQSDDIVVCNISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKK 115

Qy         75 GIQLPMDQWKILRDNIKAIDEAIKE 99
              || | :|||  ||:: : |::|: :
```

```
Db        116 GISLSVDQWNTLRNHAEDIEKALSD 140


RESULT 7
AAY44883
ID   AAY44883 standard; protein; 111 AA.
XX
AC   AAY44883;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Vernonia transcription coactivator PC4(P15) type 1.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   vernonia; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone vs1n.pk013.f21.
XX
OS   Vernonia mespilifolia.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 63
FT                   /label= Unknown
FT                   /note= "Encoded by AAN"
FT   Misc-difference 74
FT                   /label= Unknown
FT                   /note= "Encoded by GNG"
FT   Misc-difference 79
FT                   /label= Unknown
FT                   /note= "Encoded by NAC"
FT   Misc-difference 81
FT                   /label= Unknown
FT                   /note= "Encoded by AAN"
FT   Misc-difference 82
FT                   /label= Unknown
FT                   /note= "Encoded by NAA"
FT   Misc-difference 85
FT                   /label= Unknown
FT                   /note= "Encoded by GNC"
FT   Misc-difference 87
FT                   /label= Unknown
FT                   /note= "Encoded by ANA"
FT   Misc-difference 95
FT                   /label= Unknown
FT                   /note= "Encoded by CAN"
FT   Misc-difference 106
FT                   /label= Unknown
FT                   /note= "Encoded by ANT"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   N-PSDB; AAZ50416.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 6; Page 39; 43pp; English.
XX
CC   The present amino acid sequence is the vernonia Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone vs1n.pk013.f21, obtained from vs1n cDNA library, prepared from
```

```
CC   vernonia seeds. This sequence has 41% homology to PC4(P15) type 1 from
CC   Arabidopsis thaliana (gi 2997684). PC4 molecules functions as a signal
CC   mediator between activators like Gal4/VP16 and general transcription
CC   factors (GTFs) in a transcription initiation complex. The expression
CC   levels of PC4 can be manipulated and the functional properties of
CC   specific transcriptional activators can be modulated. The PC4 DNA can be
CC   used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 111 AA;

  Query Match             39.0%;  Score 206.5;  DB 3;  Length 111;
  Best Local Similarity   49.4%;  Pred. No. 3.2e-17;
  Matches   40;  Conservative   12;  Mismatches   26;  Indels    3;  Gaps    1;

Qy         17 AKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRKGI 76
              :|:  ||||   |   ||| :  :||||||:|| | |:| |||| || | |  :| : ||
Db         33 SKKDTAGDDSDDE---DDIFICDVSKNRRVSVRXWQGRVFVDIRXFYMKXGXXMPXKXGI 89

Qy         77 QLPMDQWKILRDNIKAIDEAI 97
                | || || || ::   |:|:
Db         90 SLTMDXWKELRAHVDEXDKAL 110


RESULT 8
AAG04501
ID   AAG04501 standard; protein; 85 AA.
XX
AC   AAG04501;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 551.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
```

PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.7%; Score 194.5; DB 3; Length 85;
Best Local Similarity 63.9%; Pred. No. 6.8e-16;
Matches 39; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

```
Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
              || |     ||       |||||  |||||||:|| ||||: :|||||| |||||||:|
Db         24 PAKKVAKPADD---SDQSDDIVVCNISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKK 80

Qy         75 G 75
              |
Db         81 G 81
```

RESULT 9
AAG04499
ID AAG04499 standard; protein; 120 AA.
XX
AC AAG04499;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 549.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.

```
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match           36.7%;  Score 194.5;  DB 3;  Length 120;
  Best Local Similarity   63.9%;  Pred. No. 1.1e-15;
  Matches   39;  Conservative    4;  Mismatches   15;  Indels    3;  Gaps    1;

Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
              || |     ||       |||||  |||||||:|| ||||: :|||||| ||||||||:|
Db         59 PAKKVAKPADD---SDQSDDIVVCNISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKK 115

Qy         75 G 75
              |
Db        116 G 116


RESULT 10
ABR57409
ID   ABR57409 standard; protein; 165 AA.
XX
AC   ABR57409;
XX
DT   23-OCT-2003  (revised)
DT   12-SEP-2003  (first entry)
XX
DE   Brassica campestris MIP102 protein SEQ ID NO:2.
XX
KW   Brassica campestris; plant; MIP102; virus resistance;
KW   plant virus transport protein binding protein.
XX
OS   Brassica rapa.
XX
PN   WO2003022039-A1.
XX
PD   20-MAR-2003.
XX
PF   10-SEP-2001; 2001WO-JP007858.
XX
PR   10-SEP-2001; 2001WO-JP007858.
XX
PA   (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PI   Nishiguchi M,  Nyunoya H,  Matsushita Y;
XX
DR   WPI; 2003-313169/30.
DR   N-PSDB; ACF03526.
XX
PT   Plant protein binding to plant virus transport protein for imparting
PT   virus resistance to plants including tobacco.
XX
PS   Claim 2; Fig 2; 84pp; Japanese.
XX
CC   The present invention describes a method for imparting virus resistance
CC   to plants in which the plant cells are transformed with a polynucleotide
CC   encoding a protein (I) which binds to plant virus transport protein. Also
CC   described are plants transformed by (I). The method is useful for
CC   increasing resistance of plants such as tobacco, tomato, petunia, pear,
CC   rice, wheat, barley, maize, soybean, oilseed rape, rose, apple, alfalfa,
CC   melon, lavender, onion, spinach and parsley; to viruses including
CC   tobamovirus, tobravirus, dianthovirus, alfamovirus, bromovirus,
CC   cucumovirus, comovirus, nepovirus, caulimovirus, geminivirus, potivirus
CC   and tospovirus. The present sequence represents Brassica campestris
CC   MIP102, which is a specifically claimed plant virus transport protein
CC   binding protein from the present invention. (Updated on 23-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 165 AA;

  Query Match           34.9%;  Score 185;  DB 6;  Length 165;
  Best Local Similarity   38.1%;  Pred. No. 2.5e-14;
  Matches   37;  Conservative   18;  Mismatches   22;  Indels   20;  Gaps    2;

Qy          3 RKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREF 62
              ::||| |              ||      | |::: ::|  ||| :: : || :| |||:
Db         84 KEGNKEF--------------DD------DGDLIICRLSDKRRVTIQEFRGKSLVSIREY 123

Qy         63 YEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |:|||| ||  ||| |  :||   : || ||: |:|:
Db        124 YKKDGKELPSSKGISLTDEQWSTFKKNIPAIEAAVKK 160


RESULT 11
AAY44886
ID   AAY44886 standard; protein; 141 AA.
XX
AC   AAY44886;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Soybean transcription coactivator PC4(P15) type 2.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 2;
KW   soybean; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone ses4d.pk0016.g2.
XX
OS   Glycine max.
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   N-PSDB; AAZ50419.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 18; Page 42; 43pp; English.
XX
CC   The present amino acid sequence is the soybean Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 2. It is isolated
CC   from clone ses4d.pk0016.g2, obtained from ses4d cDNA library, prepared
CC   from soybean embryogenic suspension 4 days after subculture. This
CC   sequence has 66% homology to PC4(P15) type 2 from Arabidopsis thaliana
CC   (gi 2997686). PC4 molecules functions as a signal mediator between
CC   activators like Gal4/VP16 and general transcription factors (GTFs) in a
CC   transcription initiation complex. The expression levels of PC4 can be
CC   manipulated and the functional properties of specific transcriptional
CC   activators can be modulated. The PC4 DNA can be used to create transgenic
CC   plants with altered PC4 levels, that would affect the level of
CC   transcription of specific genes in the plant. It is also used for
CC   immunological screening of cDNA libraries and to raise specific
CC   antibodies for detection. The DNA is used as probes and primers, for
CC   genetic and physical mapping of genes and as markers for traits linked to
CC   those genes
XX
SQ   Sequence 141 AA;

  Query Match           34.8%;  Score 184.5;  DB 3;  Length 141;
  Best Local Similarity   38.7%;  Pred. No. 2.3e-14;
```

```
  Matches   36;  Conservative   21;  Mismatches   25;  Indels   11;  Gaps    2;

Qy          7 KRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKD 66
              :  |||    :| :   |:|       |::: ::|  ||| :: : || :| |||:|:||
Db         56 EELGGG----SKGKEYDDEG-------DLIICRLSDKRRVTIQDFRGKTLVSIREYYKKD 104

Qy         67 GKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              || ||  ||| |  :||   : |: ||::|||:
Db        105 GKELPTSKGISLTEEQWSAFKKNVPAIEKAIKK 137
```

RESULT 12
AAG50454
ID AAG50454 standard; protein; 144 AA.
XX
AC AAG50454;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63942.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

```
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             34.5%;  Score 183;  DB 3;  Length 144;
  Best Local Similarity   38.5%;  Pred. No. 3.7e-14;
  Matches   37;  Conservative   17;  Mismatches   22;  Indels   20;  Gaps    2;

Qy          4 KGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY 63
              |||| |              ||      | |::: ::|  ||| :: : || :| |||:|
Db         64 KGNKEF--------------DD------DGDLIICRLSDKRRVTIQEFKGKSLVSIREYY 103

Qy         64 EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                :|||| ||  ||| |  :||   : |: ||: |:|:
Db        104 KKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK 139


RESULT 13
AAG09966
ID   AAG09966 standard; protein; 144 AA.
XX
AC   AAG09966;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 8102.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
```

PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 183; DB 3; Length 144;
Best Local Similarity 38.5%; Pred. No. 3.7e-14;
Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;

```
Qy          4 KGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY 63
              |||| |               ||       | |::: ::|  ||| :: : || :| |||:|
Db         64 KGNKEF-------------DD------DGDLIICRLSDKRRVTIQEFKGKSLVSIREYY 103
```

```
Qy         64 EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :|||| ||  ||| |  :||   : |: ||: |:|:
Db        104 KKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK 139
```

RESULT 14
AAG50453
ID AAG50453 standard; protein; 165 AA.
XX
AC AAG50453;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63941.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 183; DB 3; Length 165;
Best Local Similarity 38.5%; Pred. No. 4.4e-14;
Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;

```
Qy          4 KGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY 63
              ||||.|               ||       |.|::: ::|  .||| :: : || :| |||:|
Db         85 KGNKEF--------------DD------DGDLIICRLSDKRRVTIQEFKGKSLVSIREYY 124

Qy         64 EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :|||| ||  ||| |  :||   : |: ||: |:|:
Db        125 KKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK 160
```

RESULT 15
AAG09965
ID AAG09965 standard; protein; 165 AA.
XX
AC AAG09965;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8101.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 183; DB 3; Length 165;
Best Local Similarity 38.5%; Pred. No. 4.4e-14;
Matches 37; Conservative 17; Mismatches 22; Indels 20; Gaps 2;

```
Qy          4 KGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY 63
              |||| |               ||       | |::: ::|  ||| :: : || :| |||:|
Db         85 KGNKEF-------------DD------DGDLIICRLSDKRRVTIQEFKGKSLVSIREYY 124

Qy         64 EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                :|||| ||  ||| |  :||   : |: ||: |:|:
Db        125 KKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK 160
```

Search completed: January 10, 2005, 23:03:29

Job time : 156 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:57:52 ; Search time 39 Seconds
(without alignments)
249.177 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	247.5	46.7	107	2	T52113	probable transcrip
2	183	34.5	165	2	T52114	probable transcrip
3	172	32.5	127	2	A28084	DNA-binding protei
4	166	31.3	119	2	A23063	pancreatic B-cell
5	160.5	30.3	127	2	A54670	RNA polymerase II
6	133.5	25.2	136	2	T37767	probable transcrip
7	117	22.1	292	2	S52888	SUB1 protein - yea
8	112	21.1	98	2	S66164	single strand DNA-
9	102	19.2	124	2	T24876	hypothetical prote
10	87	16.4	421	2	T47393	hypothetical prote
11	81	15.3	342	2	D49348	succinoglycan bios
12	81	15.3	342	2	G95975	glucosyltransferas
13	77.5	14.6	288	2	S40173	ExoU protein - Rhi
14	76	14.3	160	2	AI1680	dihydrofolate redu
15	74.5	14.1	177	2	D84828	AP2 domain transcr
16	74	14.0	241	2	F97202	hypothetical prote
17	73.5	13.9	830	2	A84951	DNA topoisomerase
18	73	13.8	492	2	A87471	hypothetical prote
19	72.5	13.7	483	2	A53918	chitinase (EC 3.2.
20	72	13.6	160	2	AI1308	dihydrofolate redu
21	71.5	13.5	67	2	F89778	hypothetical prote
22	71	13.4	127	2	T39422	hypothetical prote
23	71	13.4	263	2	T11556	nef protein - simi
24	70.5	13.3	128	2	S13069	ribosomal protein
25	70.5	13.3	159	2	AC0513	dihydrofolate redu
26	70.5	13.3	159	2	B39799	dihydrofolate redu
27	70	13.2	566	2	T46933	hypothetical prote
28	70	13.2	630	1	XXECDP	dihydrolipoamide S
29	70	13.2	630	2	G90643	hypothetical prote
30	70	13.2	630	2	G85494	hypothetical prote
31	69.5	13.1	239	2	S54852	nef protein - simi
32	69.5	13.1	244	2	S61208	NEF protein - simi
33	69.5	13.1	250	2	S61207	NEF protein - simi
34	69.5	13.1	1096	2	T21091	hypothetical prote
35	69	13.0	490	2	A87672	tldD protein [impo
36	68.5	12.9	434	2	G70422	tryptophan synthas
37	68.5	12.9	505	2	G70954	probable lysS prot
38	68.5	12.9	2295	2	B71621	probable membrane
39	68	12.8	126	2	T06861	ribosomal protein
40	68	12.8	163	2	B84081	dihydrofolate redu
41	68	12.8	246	2	S54853	nef protein - simi
42	68	12.8	251	2	S54867	nef protein - simi
43	68	12.8	285	2	F69375	ABC transporter, A
44	68	12.8	449	2	T05242	hypothetical prote
45	68	12.8	454	2	T36209	probable pyridine

ALIGNMENTS

RESULT 1
T52113
probable transcription co-activator KIWI [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52113
R;Cormack, R.S.; Hahlbrock, K.; Somssich, I.E.
Plant J. 14, 685-92, 1998
A;Title: Isolation of putative plant transcriptional coactivators using a modified two-h
A;Reference number: Z25848; MUID:98346011; PMID:9681033
A;Accession: T52113
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-107 <COR>
A;Cross-references: UNIPROT:O65154; EMBL:AF053302; PIDN:AAC08574.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: KIWI
A;Map position: V
C;Function:
A;Description: probably plays a role in gene activation during pathogen defence and plant
C;Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2

Query Match 46.7%; Score 247.5; DB 2; Length 107;
Best Local Similarity 56.5%; Pred. No. 3.1e-17;
Matches 48; Conservative 12; Mismatches 22; Indels 3; Gaps 1;

```
Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
              || |     ||       |||||  |||||||:|| |||| : |||||| ||||||||:|
Db         24 PAKKVAKPADD---SDQSDDIVVCNISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKK 80

Qy         75 GIQLPMDQWKILRDNIKAIDEAIKE 99
                 || | :|||  ||:: : |::|: :
Db         81 GISLSVDQWNTLRNHAEDIEKALSD 105
```

RESULT 2
T52114
probable transcription co-activator KELP [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F25I24.130
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52114; T01885; T04287
R;Cormack, R.S.; Hahlbrock, K.; Somssich, I.E.
Plant J. 14, 685-92, 1998
A;Title: Isolation of putative plant transcriptional coactivators using a modified two-h
A;Reference number: Z25848; MUID:98346011; PMID:9681033
A;Accession: T52114
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-165 <COR>
A;Cross-references: UNIPROT:O65155; EMBL:AF053303; PIDN:AAC08575.1

```
A;Experimental source: cultivar Columbia
R;Madsen, C.; Graves, T.; Cotton, M.; Modde, T.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of A. thaliana F8M12.
A;Reference number: Z14450
A;Accession: T01885
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-165 <MAD>
A;Cross-references: EMBL:AF080118; NID:g3513725; PID:g3513735
A;Experimental source: cultivar Columbia
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15261
A;Accession: T04287
A;Molecule type: DNA
A;Residues: 1-165 <BEV>
A;Cross-references: EMBL:AL049525
A;Experimental source: cultivar Columbia; BAC clone F25I24
C;Genetics:
A;Gene: KELP
A;Map position: IV
A;Introns: 100/3; 136/1
A;Note: F8M12.6; F25I24.130
C;Function:
A;Description: probably plays a role in gene activation during pathogen defence and plan

  Query Match             34.5%;  Score 183;  DB 2;  Length 165;
  Best Local Similarity   38.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   17;  Mismatches   22;  Indels   20;  Gaps    2;

Qy          4 KGNKRPGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY 63
              |||| |              ||      | |::: ::|  ||| :: : || :| |||:|
Db         85 KGNKEP--------------DD------DGDLIICRLSDKRRVTIQEFKGKSLVSIREYY 124

Qy         64 EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :|||| ||  ||| |  :||   : |: ||: |:|:
Db        125 KKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK 160


RESULT 3
A28084
DNA-binding protein p9 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A28084
R;Ballard, D.W.; Philbrick, W.M.; Bothwell, A.L.M.
J. Biol. Chem. 263, 8450-8457, 1988
A;Title: Identification of a novel 9-kDa polypeptide from nuclear extracts. DNA binding
A;Reference number: A28084; MUID:88228080; PMID:3372536
A;Accession: A28084
A;Molecule type: mRNA
A;Residues: 1-127 <BAL>
A;Cross-references: UNIPROT:P11031; GB:J03750; NID:g192235; PIDN:AAA37317.1; PID:g309120
C;Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2
C;Keywords: DNA binding

  Query Match             32.5%;  Score 172;  DB 2;  Length 127;
  Best Local Similarity   41.4%;  Pred. No. 9e-10;
  Matches   36;  Conservative   20;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
              :|    ||       | |: || :  || | | |:|| : ||:::||||:: : :|:  ||
Db         41 KPGETSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPG 99

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |||| | |:||  |:: |  ||:|:::
Db        100 RKGISLNMEQWSQLKEQISDIDDAVRK 126


RESULT 4
A23063
pancreatic B-cell tumor protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Aug-1987 #sequence_revision 08-Aug-1987 #text_change 09-Jul-2004
C;Accession: A23063
R;Soma, G.I.; Kitahara, N.; Andoh, T.
Biochem. Biophys. Res. Commun. 124, 164-171, 1984
A;Title: Molecular cloning and characterization of a cDNA clone for a protein specificall
A;Reference number: A23063; MUID:85046489; PMID:6208900
A;Accession: A23063
A;Molecule type: mRNA
A;Residues: 1-119 <SOM>
A;Cross-references: UNIPROT:Q63396; GB:K02816; NID:g205853; PIDN:AAA41758.1; PID:g205854
C;Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2

  Query Match             31.3%;  Score 166;  DB 2;  Length 119;
  Best Local Similarity   40.2%;  Pred. No. 3.2e-09;
  Matches   35;  Conservative   21;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
              :|:   ||       | |: || :  || | | |:|| : ||:::||||:: : :|:  |
Db         33 KPSESSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPR 91

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |||| | |:||  |:: |  ||:|:::
Db         92 RKGISLNMEQWSQLKEQISDIDDAVRK 118


RESULT 5
A54670
RNA polymerase II transcription cofactor p15 - human
N;Alternate names: DNA binding protein PC4
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: A54670; A54669
R;Kretzschmar, M.; Kaiser, K.; Lottspeich, F.; Meisterernst, M.
Cell 78, 525-534, 1994
A;Title: A novel mediator of class II gene transcription with homology to viral immediate
A;Reference number: A54670; MUID:94340741; PMID:8062392
A;Accession: A54670
A;Molecule type: mRNA
A;Residues: 1-127 <KRE>
A;Cross-references: UNIPROT:P53999; GB:X79805; NID:g619160; PIDN:CAA56200.1; PID:g619161
R;Ge, H.; Roeder, R.G.
Cell 78, 513-523, 1994
A;Title: Purification, cloning, and characterization of a human coactivator, PC4, that me
A;Reference number: A54669; MUID:94340740; PMID:8062391
A;Accession: A54669
A;Molecule type: mRNA
A;Residues: 1-127 <GEA>
A;Cross-references: GB:U12979; NID:g531394; PIDN:AAA20980.1; PID:g531395
A;Note: parts of this sequence, including the amino end of the mature protein, were confi
C;Genetics:
A;Gene: GDB:P15
A;Cross-references: GDB:453302
C;Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2
C;Keywords: DNA binding; nucleus; phosphoprotein
F;2-127/Product: RNA polymerase II transcription cofactor p15 #status experimental <MAT>

  Query Match             30.3%;  Score 160.5;  DB 2;  Length 127;
  Best Local Similarity   37.9%;  Pred. No. 1.2e-08;
  Matches   36;  Conservative   21;  Mismatches   27;  Indels   11;  Gaps    3;

Qy         14 EPAAKRRAAGDDG--------PSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-E 64
              |   |::  |:           | | ||::   || | | |:|| : |||::||||:: :
Db         34 EKPVKKQKTGETSRALSSSKQSSSSRDDNMF--QIGKMRYVSVRDFKGKVLIDIREYWMD 91

Qy         65 KDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :|:  |||||| |  :||  |:: |  ||:|:::
Db         92 PEGEMKPGRKGISLNPEQWSQLKEQISDIDDAVRK 126


RESULT 6
T37767
probable transcription activator - fission yeast (Schizosaccharomyces pombe)
```

```
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37767
R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z21738
A;Accession: T37767
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-136 <MUR>
A;Cross-references: UNIPROT:P87294; EMBL:Z97185; PIDN:CAB10003.1; GSPDB:GN00066; SPDB:SP
A;Experimental source: strain 972h-; cosmid c16A10
C;Genetics:
A;Gene: SPDB:SPAC16A10.02
A;Map position: 1
A;Introns: 31/3

  Query Match             25.2%;  Score 133.5;  DB 2;  Length 136;
  Best Local Similarity   31.5%;  Pred. No. 5.6e-06;
  Matches   29;  Conservative   25;  Mismatches   33;  Indels    5;  Gaps    2;

Qy          7 KRFGGGGEPAAKRRAAGDDGPSESADDDIVVA-QISKNRRVAVRTWNGKVVVDIREFYEK 65
              ||    |: |: ::   :    : :| ::  |   :: :|: :  : |   | |||:|||
Db          4 KRVIASGDKASSKKPKTE----KQSDHELHWALNETEKKRITLSEFRGTRYVHIREYYEK 59

Qy         66 DGKTLPGRKGIQLPMDQWKILRDNIKAIDEAI 97
              ||  |||:||| | :::|| |:  |  :|:::
Db         60 DGDMLPGKKGIALNINEWKKLKQLIHEVDDSL 91


RESULT 7
S52888
SUB1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YM9532.04c; protein YMR039c; suppressor protein SUB1
C;Species: Saccharomyces cerevisiae
C;Date: 08-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S52888; S69242; S69243
R;Odell, C.; Bowman, S.
submitted to the EMBL Data Library, February 1995
A;Reference number: S52885
A;Accession: S52888
A;Molecule type: DNA
A;Residues: 1-292 <ODE>
A;Cross-references: UNIPROT:P54000; EMBL:Z48502; NID:g695715; PID:g695719; MIPS:YMR039c
R;Knaus, R.; Pollock, R.; Guarente, L.
EMBO J. 15, 1933-1940, 1996
A;Title: Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-
A;Reference number: S69242; MUID:96203115; PMID:8617240
A;Accession: S69242
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-169,'E',171-233,'E',235-292 <KNA>
A;Accession: S69243
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-169,'E',171-233,'E',235-292 <KNW>
C;Genetics:
A;Gene: SGD:SUB1
A;Cross-references: SGD:S0004642; MIPS:YMR039c
A;Map position: 13R
C;Function:
A;Description: SUB1 protein plays a role in the release of transcription factor IIB from
C;Keywords: DNA binding; nucleus

  Query Match             22.1%;  Score 117;  DB 2;  Length 292;
  Best Local Similarity   32.3%;  Pred. No. 0.00054;
  Matches   32;  Conservative   16;  Mismatches   47;  Indels    4;  Gaps    2;

Qy          2 WRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIRE 61
              :|   :   |||  :      |      || |   :  : ||:|| || :    ::||||
Db          7 YRNKRRSDNGGGNLSNSNNNNGGMPSGLSASD--AIFDLGKNKRVTVRQFRNINLIDIRE 64

Qy         62 FY--EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIK 98
                :|     |:  ||:||| |  | :  |  :   ||||::
Db         65 YYLDSSTGEMKPGKKGISLTEDLYDELLKHRLNIDEALR 103


RESULT 8
S66164
single strand DNA-binding protein C31A - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66164
R;Lacoste, J.; Codani-Simonart, S.; Best-Belpomme, M.; Peronnet, F.
Nucleic Acids Res. 23, 5073-5079, 1995
A;Title: Characterization and cloning of p11, a transrepressor of Drosophila melanogaste
A;Reference number: S66164; MUID:96145266; PMID:8559667
A;Accession: S66164
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-98 <LAC>
A;Cross-references: UNIPROT:Q24026; EMBL:U18971; NID:g1109661; PIDN:AAC46989.1; PID:g110

  Query Match             21.1%;  Score 112;  DB 2;  Length 98;
  Best Local Similarity   35.7%;  Pred. No. 0.00049;
  Matches   25;  Conservative   15;  Mismatches   26;  Indels    4;  Gaps    2;

Qy         14 EPAAKRRAAGD---DGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTL 70
              :||:|:    |     | :| ::      :   |:| :  : |:  |||||||:| |: |
Db         22 KPASKKAKESDAPNSDPKDSGENGATSWTLEGLRQVRINEFRGRKSVDIREFYDKGGQIL 81

Qy         71 PGRKGIQLPM 80
              ||:||  ||:
Db         82 PGKKG-HLPI 90


RESULT 9
T24876
hypothetical protein T13F2.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24876
R;Swinburne, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19947
A;Accession: T24876
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-124 <WIL>
A;Cross-references: UNIPROT:Q94045; EMBL:Z81122; PIDN:CAB03353.1; GSPDB:GN00022; CESP:T1
A;Experimental source: clone T13F2
C;Genetics:
A;Gene: CESP:T13F2.2
A;Map position: 4
A;Introns: 27/3; 61/3; 91/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2

  Query Match             19.2%;  Score 102;  DB 2;  Length 124;
  Best Local Similarity   42.4%;  Pred. No. 0.0061;
  Matches   25;  Conservative    9;  Mismatches   23;  Indels    2;  Gaps    2;

Qy         39 QISKNRRVAVRTWNGKVVVDIREFY-EKDG-KTLPGRKGIQLPMDQWKILRDNIKAIDE 95
              :|   |   |  : ||  |:|||:| ::|  | :| |||| |   ||  |:| |  ||:
Db         64 EIGNLRYATVSKFKGKEYVNIREYYIDRDSQKMMPSRKGISLSKAQWANLKDLIPEIDK 122


RESULT 10
T47393
hypothetical protein T18D12.80 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47393
R;Massenet, O.; Clabault, G.; Quigley, F.; Mache, R.; Mewes, H.W.; Lemcke, K.; Mayer, K.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24464
A;Accession: T47393
```

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-421 <MAS>
A;Cross-references: UNIPROT:Q9M239; EMBL:AL138644
A;Experimental source: cultivar Columbia; BAC clone T18D12
C;Genetics:
A;Map position: 3
A;Introns: 66/3; 101/1; 143/2
A;Note: T18D12.80

  Query Match             16.4%;  Score 87;  DB 2;  Length 421;
  Best Local Similarity   27.3%;  Pred. No. 0.69;
  Matches   27;  Conservative   16;  Mismatches   34;  Indels   22;  Gaps    3;

Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY---EKDGK--- 68
              || ||: :        |:::  : : |: : | |  : ||  : | : :   ||| |
Db         49 PAYKRKVS-------VAEEERFLYKFSETKDVMVNKYQGKTSLLIGKKHVRKEKDAKFRA 101

Qy         69 ---------TLPGRKGIQLPMDQWKILRDNIKAIDEAIK 98
                      |    ||  |  || : : |  | :||||
Db        102 TLVNDEKAFMLCSLLGISFPTHQWSVFKKNFSATEEAIK 140


RESULT 11
D49348
succinoglycan biosynthesis glycosyltransferase (EC 2.4.1.-) exoU [validated] - Rhizobium
C;Species: Rhizobium meliloti
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: D49348
R;Glucksmann, M.A.; Reuber, T.L.; Walker, G.C.
J. Bacteriol. 175, 7033-7044, 1993
A;Title: Family of glycosyl transferases needed for the synthesis of succinoglycan by Rh
A;Reference number: A49348; MUID:94042869; PMID:8226645
A;Accession: D49348
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-342 <GLU>
A;Cross-references: UNIPROT:P33700; GB:L20758; NID:g393240; PIDN:AAA16053.1; PID:g393252
C;Genetics:
A;Gene: exoU
C;Function:
A;Description: EC 2.4.1.-; succinoglycan biosynthesis glycosyltransferase [validated, MU
, exoU, and exoW together are responsible for sugar addition to the lipid carrier
A;Pathway: succinoglycan biosynthesis
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             15.3%;  Score 81;  DB 2;  Length 342;
  Best Local Similarity   27.9%;  Pred. No. 2.1;
  Matches   29;  Conservative    9;  Mismatches   36;  Indels   30;  Gaps    3;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQ--------------------ISKNRRVAVRTWN 52
              || |       |||   | ||   ||:                        ::|  :|:
Db         32 EPEAAEVVVIDDG---STDDSASVARAADDGTGRLNVVRFEENRGPAAARNHAIAISHSP 88

Qy         53 GKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEA 96
                 |:|  :|:       ||| |  |  | |  : |||  || |
Db         89 LIGVLDADDFF------FPGRLGQLLSQDGWDFIADNIAFIDAA 126


RESULT 12
G95975
glucosyltransferase protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmi
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: G95975
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: G95975
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-342 <KUR>
A;Cross-references: UNIPROT:P33700; GB:AL591985; PIDN:CAC49471.1; PID:g15140957; GSPDB:GN
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: exoU; SMb20948
A;Genome: plasmid

  Query Match             15.3%;  Score 81;  DB 2;  Length 342;
  Best Local Similarity   27.9%;  Pred. No. 2.1;
  Matches   29;  Conservative    9;  Mismatches   36;  Indels   30;  Gaps    3;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQ--------------------ISKNRRVAVRTWN 52
              || |       |||   | ||   ||:                        ::|  :|:
Db         32 EPEAAEVVVIDDG---STDDSASVARAADDGTGRLNVVRFEENRGPAAARNHAIAISHSP 88

Qy         53 GKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEA 96
                 |:|  :|:       ||| |  |  | |  : |||  || |
Db         89 LIGVLDADDFF------FPGRLGQLLSQDGWDFIADNIAFIDAA 126


RESULT 13
S40173
ExoU protein - Rhizobium meliloti
C;Species: Rhizobium meliloti
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: S40173
R;Becker, A.; Kleickmann, A.; Kuester, H.; Keller, M.; Arnold, W.; Puehler, A.
submitted to the EMBL Data Library, April 1993
A;Description: Analysis of the Rhizobium meliloti genes exoU, exoV, exoW, exoT and exoI i
ses.
A;Reference number: S40173
A;Accession: S40173
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-288 <BEC>
A;Cross-references: EMBL:Z22646

  Query Match             14.6%;  Score 77.5;  DB 2;  Length 288;
  Best Local Similarity   29.0%;  Pred. No. 3.9;
  Matches   29;  Conservative   12;  Mismatches   36;  Indels   23;  Gaps    4;

Qy         14 EPAAKRRAAGDDGPSE-------SADDD---IVVAQISKNRR-------VAVRTWNGKVV 56
              || |       ||| ::       :|||    : | :  :||       :|:       |
Db         32 EPEAAEVVVIDDGSTDDSASVARAADDGTGRLNVVRFEENRGPARAQSCIAISHSPLIGV 91

Qy         57 VDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEA 96
              :|  :|:       ||| |  |  | |  : |||  || |
Db         92 LDADDFF------FPGRLGQLLSQDGWDFIADNIAFIDAA 125


RESULT 14
AI1680
dihydrofolate reductases homolog lin1987 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C;Accession: AI1680
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1680
```

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-160 <GLA>
A;Cross-references: UNIPROT:Q92AD5; GB:AL592022; PIDN:CAC97217.1; PID:g16414488; GSPDB:G
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1987
C;Superfamily: Dihydrofolate reductase; type I dihydrofolate reductase homology

  Query Match             14.3%;  Score 76;  DB 2;  Length 160;
  Best Local Similarity   40.0%;  Pred. No. 2.8;
  Matches   22;  Conservative    8;  Mismatches   17;  Indels    8;  Gaps    2;

Qy         49 RTWNGKVVVDIREFYEKDGKTLPGRKGI------QLPMDQWKIL--RDNIKAIDE 95
              :|  || :|  |: ||  || || || |      :| :|  :||  || : |: |
Db         33 KTTTGKTLVMGRKTYESLGKALPNRKTIVLTRDNELQLDDAEILHSRDEVLALAE 87


RESULT 15
D84828
AP2 domain transcription factor [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84828
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84828
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-177 <STO>
A;Cross-references: UNIPROT:Q9SIZ0; GB:AE002093; NID:g4586052; PIDN:AAD25670.1; GSPDB:GN
C;Genetics:
A;Gene: At2g40350
A;Map position: 2

  Query Match             14.1%;  Score 74.5;  DB 2;  Length 177;
  Best Local Similarity   30.4%;  Pred. No. 4.4;
  Matches   34;  Conservative   11;  Mismatches   32;  Indels   35;  Gaps    7;

Qy          2 WRKGNKRF-------GGGGEPAAK---RRA-----AGDDGPSESADDDIVVAQISKNRRV 46
              ||: |::        ||| :|  |   :|:      |  ||     |   | |
Db         20 WREYNEQTEADSCIDGGGSKPIRKAPPKRSRKGCMKGKGGPENGICDYTGVRQ------- 72

Qy         47 AVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIK 98
                ||| || | :|||         ||| | :| :  :    :   | ||| |
Db         73 --RTW-GKWVAEIRE---------PGR-GAKLWLGTFSSSYEAALAYDEASK 111


Search completed: January 10, 2005, 23:07:31
Job time : 41 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:49:56 ; Search time 193 Seconds
(without alignments)
301.103 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	530	100.0	101	2	Q6Z777	Q6z777 oryza sativ
2	530	100.0	101	2	BAD15868	Bad15868 oryza sat
3	247.5	46.7	107	1	KIWI_ARATH	O65154 arabidopsis
4	185	34.9	165	2	Q9AVE8	Q9ave8 brassica ca
5	183	34.5	165	1	KELP_ARATH	O65155 arabidopsis
6	180.5	34.1	184	2	Q94JE5	Q94je5 oryza sativ
7	172	32.5	126	1	TCP4_MOUSE	P11031 mus musculu
8	172	32.5	127	2	BAC33314	Bac33314 mus muscu
9	172	32.5	127	2	BAC33642	Bac33642 mus muscu
10	166	31.3	119	1	TCP4_RAT	Q63396 rattus norv
11	165	31.1	110	2	Q9FY90	Q9fy90 arabidopsis
12	161.5	30.5	73	2	Q6E433	Q6e433 homo sapien
13	160.5	30.3	126	1	TCP4_HUMAN	P53999 homo sapien
14	160.5	30.3	127	2	CAG33183	Cag33183 homo sapi
15	160	30.2	124	2	Q6DGP6	Q6dgp6 brachydanio
16	153	28.9	128	2	Q6NTX9	Q6ntx9 xenopus lae
17	153	28.9	128	2	AAH68824	Aah68824 xenopus l
18	151	28.5	128	2	Q6DJR3	Q6djr3 xenopus tro
19	149	28.1	136	2	Q6C485	Q6c485 yarrowia li
20	145.5	27.5	110	1	TCP4_DROME	Q9vlr5 drosophila
21	137	25.8	594	2	Q93YB6	Q93yb6 nicotiana t
22	133.5	25.2	136	1	TCP4_SCHPO	P87294 schizosacch
23	133.5	25.2	273	2	Q75DD4	Q75dd4 ashbya goss
24	133.5	25.2	273	2	AAS50863	Aas50863 ashbya go
25	118	22.3	172	1	TCP4_NEUCR	Q872f4 neurospora
26	117	22.1	135	2	Q6BPT2	Q6bpt2 debaryomyce
27	117	22.1	292	1	SUB1_YEAST	P54000 saccharomyc
28	112	21.1	98	2	Q24026	Q24026 drosophila
29	112	21.1	277	2	Q6FU60	Q6fu60 candida gla
30	102	19.2	124	1	TCP4_CAEEL	Q94045 caenorhabdi
31	95	17.9	244	2	Q6CIG4	Q6cig4 kluyveromyc
32	95	17.9	409	2	Q6DBG5	Q6dbg5 arabidopsis
33	88	16.6	269	1	FPG_VIBPA	Q87t81 vibrio para
34	87	16.4	421	2	Q9M239	Q9m239 arabidopsis
35	86.5	16.3	488	2	Q8GXX5	Q8gxx5 arabidopsis
36	81	15.3	71	2	Q72LS9	Q72ls9 leptospira
37	81	15.3	71	2	Q8CXR1	Q8cxr1 leptospira
38	81	15.3	71	2	AAS72010	Aas72010 leptospir
39	81	15.3	342	1	EXOU_RHIME	P33700 rhizobium m
40	81	15.3	1018	2	Q6L0V3	Q6l0v3 picrophilus
41	80.5	15.2	269	1	FPG_VIBVY	Q7mps3 vibrio vuln
42	80	15.1	3985	2	Q7Q293	Q7q293 anopheles g
43	78.5	14.8	269	1	FPG_VIBVU	Q8ddy4 vibrio vuln
44	78.5	14.8	1646	2	Q7PRH5	Q7prh5 anopheles g
45	76.5	14.4	271	2	Q8HYL3	Q8hyl3 bos taurus

ALIGNMENTS

RESULT 1
Q6Z777
ID Q6Z777 PRELIMINARY; PRT; 101 AA.
AC Q6Z777;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional coactivator p15 (PC4) family protein-like.
GN Name=P0470G10.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP004876; BAD15868.1; -.
DR InterPro; IPR003173; PC4.
DR InterPro; IPR009044; ssDNA_bind_regul.
DR Pfam; PF02229; PC4; 1.
SQ SEQUENCE 101 AA; 11278 MW; 43AA93C0B4B88C92 CRC64;

Query Match 100.0%; Score 530; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e-45;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60

Qy         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
              |||||||||||||||||||||||||||||||||||||||||
Db         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
```

RESULT 2
BAD15868
ID BAD15868 PRELIMINARY; PRT; 101 AA.
AC BAD15868;
DT 23-APR-2004 (TrEMBLrel. 27, Created)
DT 23-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 23-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional coactivator p15 (PC4) family protein-like.
GN P0470G10.10.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;

```
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
RT   clone:P0470G10.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004876; BAD15868.1; -.
SQ   SEQUENCE   101 AA;  11278 MW;  43AA93C0B4B88C92 CRC64;

  Query Match             100.0%;  Score 530;  DB 2;  Length 101;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-45;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60

Qy         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
              |||||||||||||||||||||||||||||||||||||||||
Db         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101


RESULT 3
KIWI_ARATH
ID   KIWI_ARATH     STANDARD;      PRT;   107 AA.
AC   O65154;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   RNA polymerase II transcriptional coactivator KIWI.
GN   Name=KIWI; OrderedLocusNames=At5g09250; ORFNames=T5E8_50;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98346011; PubMed=9681033;
RA   Cormack R.S., Hahlbrock K., Somssich I.E.;
RT   "Isolation of putative plant transcriptional coactivators using a
RT   modified two-hybrid system incorporating a GFP reporter gene.";
RL   Plant J. 14:685-692(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016721; PubMed=11130714; DOI=10.1038/35048507;
RA   Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA   Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA   Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA   Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA   Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA   Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA   Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA   Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA   Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA   Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,
RA   Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,
RA   Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA   Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
RA   Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA   Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA   Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA   Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA   Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA   van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
RA   Weitzenegger T., Bothe G., Rose M., Hauf J., Berneiser S., Hempel S.,
RA   Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA   Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA   Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
RT   "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:823-826(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
RL   Science 302:842-846(2003).
CC   -!- FUNCTION: General coactivator that functions cooperatively with
CC       TAFs and mediates functional interactions between upstream
CC       activators and the general transcriptional machinery. Binds
CC       single-stranded DNA (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: Belongs to the transcriptional coactivator PC4 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF053302; AAC08574.1; -.
DR   EMBL; AL391712; CAC05451.1; -.
DR   EMBL; BT004745; AAO44011.1; -.
DR   PIR; T52113; T52113.
DR   HSSP; P53999; 1PCF.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
KW   Activator; DNA-binding; Nuclear protein; Transcription;
KW   Transcription regulation.
SQ   SEQUENCE   107 AA;  12083 MW;  513F84BCCB240486 CRC64;

  Query Match             46.7%;  Score 247.5;  DB 1;  Length 107;
  Best Local Similarity   56.5%;  Pred. No. 5.8e-17;
  Matches   48;  Conservative   12;  Mismatches   22;  Indels    3;  Gaps    1;

Qy         15 PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
              || |      ||       |||||  |||||||:|| |||| : :|||||| ||||||||:|
Db         24 PAKKVAKPADD---SDQSDDIVVCNISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKK 80

Qy         75 GIQLPMDQWKILRDNIKAIDEAIKE 99
              || | :|||  ||:: : |::|: :
Db         81 GISLSVDQWNTLRNHAEDIEKALSD 105


RESULT 4
Q9AVE8
ID   Q9AVE8      PRELIMINARY;      PRT;   165 AA.
AC   Q9AVE8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative transcriptional coactivator.
GN   Name=BcKELP;
OS   Brassica campestris (Field mustard).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3711;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S9 homozygote;
RX   MEDLINE=21445625; PubMed=11561731;
RA   Matsushita Y., Deguchi M., Youda M., Nishiguchi M., Nyunoya H.;
RT   "The tomato mosaic tobamovirus movement protein interacts with a
RT   putative transcriptional coactivator KELP.";
RL   Mol. Cells 12:57-66(2001).
DR   EMBL; AB050390; BAB41214.1; -.
DR   HSSP; P53999; 1PCF.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0003713; F:transcription coactivator activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
SQ   SEQUENCE   165 AA;  19239 MW;  FB3A0BB5CDFD3384 CRC64;

  Query Match             34.9%;  Score 185;  DB 2;  Length 165;
  Best Local Similarity   38.1%;  Pred. No. 1.7e-10;
  Matches   37;  Conservative   18;  Mismatches   22;  Indels   20;  Gaps    2;

Qy          3 RKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREF 62
              ::||| |               ||       | |::: ::|  ||| :: : || :| |||:
Db         84 KEGNKEF--------------DD------DGDLIICRLSDKRRVTIQEFRGKSLVSIREY 123

Qy         63 YEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                |:|||| ||  ||| |  :||   : || ||: |:|:
Db        124 YKKDGKELPSSKGISLTDEQWSTFKKNIPAIEAAVKK 160


RESULT 5
KELP_ARATH
ID   KELP_ARATH     STANDARD;      PRT;   165 AA.
AC   O65155;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   RNA polymerase II transcriptional coactivator KELP.
GN   Name=KELP; OrderedLocusNames=At4g10920; ORFNames=F8M12.6, F25I24.130;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98346011; PubMed=9681033;
RA   Cormack R.S., Hahlbrock K., Somssich I.E.;
RT   "Isolation of putative plant transcriptional coactivators using a
RT   modified two-hybrid system incorporating a GFP reporter gene.";
RL   Plant J. 14:685-692(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
RL   Science 302:842-846(2003).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.A.;
RT   "Full-length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: General coactivator that functions cooperatively with
CC       TAFs and mediates functional interactions between upstream
CC       activators and the general transcriptional machinery. Binds
CC       single-stranded DNA (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: Belongs to the transcriptional coactivator PC4 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF053303; AAC08575.1; -.
DR   EMBL; AF080118; AAC33951.1; -.
DR   EMBL; AL049525; CAB40060.1; -.
```

```
DR   EMBL; AL161518; CAB81193.1; -.
DR   EMBL; AY085896; AAM63108.1; -.
DR   EMBL; BT005847; AAO64782.1; -.
DR   PIR; T52114; T52114.
DR   HSSP; P53999; 1PCF.
DR   GeneFarm; 2402; -.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
KW   Activator; DNA-binding; Nuclear protein; Transcription;
KW   Transcription regulation.
SQ   SEQUENCE   165 AA;  19177 MW;  6620F1693D40D3D1 CRC64;

  Query Match             34.5%;  Score 183;  DB 1;  Length 165;
  Best Local Similarity   38.5%;  Pred. No. 2.7e-10;
  Matches   37;  Conservative   17;  Mismatches   22;  Indels   20;  Gaps    2;

Qy          4 KGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY 63
              |||| |              ||      | |::: ::|  ||| :: : || :| |||:|
Db         85 KGNKEF-------------DD------DGDLIICRLSDKRRVTIQEFKGKSLVSIREYY 124

Qy         64 EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :|||| ||  ||| |  :||   : |: ||: |:|:
Db        125 KKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK 160


RESULT 6
Q94JE5
ID   Q94JE5      PRELIMINARY;      PRT;   184 AA.
AC   Q94JE5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   P0684B02.23 protein (P0692C11.10 protein).
GN   Name=P0684B02.23; Synonyms=P0692C11.10;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA   Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA   Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA   Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA   Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA   Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA   Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA   Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA   Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA   Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,
RA   Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA   Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA   Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA   Yano M., Jiang J., Gojobori T.;
RT   "The genome sequence and structure of rice chromosome 1.";
RL   Nature 420:312-316(2002).
DR   EMBL; AP003023; BAB44036.1; -.
DR   EMBL; AP003381; BAB86514.1; -.
DR   HSSP; P53999; 1PCF.
DR   Gramene; Q94JE5; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0003713; F:transcription coactivator activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
SQ   SEQUENCE   184 AA;  21028 MW;  BE1EFAD4C6340B5A CRC64;

  Query Match             34.1%;  Score 180.5;  DB 2;  Length 184;
  Best Local Similarity   36.5%;  Pred. No. 5.4e-10;
  Matches   38;  Conservative   22;  Mismatches   29;  Indels   15;  Gaps    3;

Qy         10 GGGGEPAAKRRAAGDD-----------GPSESADD--DIVVAQISKNRRVAVRTWNGKV 55
              || ||     :|  ::           |      ||  |::: ::|  ||| :: : ||
Db         76 GGAGEEGKDEQAEEEEEEEEEEEERGGGTKREYDDEGDLILCRLSARRRVTLQEFKGKT 135

Qy         56 VVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :: |||:| |||| ||  ||| | ::||:  ||:: ||::|||:
Db        136 LLSIREYYFKDGKELPA-KGISLTVEQWEAFRDSVPAIEDAIKK 178


RESULT 7
TCP4_MOUSE
ID   TCP4_MOUSE     STANDARD;      PRT;   126 AA.
AC   P11031;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Activated RNA polymerase II transcriptional coactivator p15 precursor
DE   (Positive cofactor 4) (PC4) (p14) (Single-stranded DNA binding protein
DE   p9).
GN   Name=Rpo2tc1; Synonyms=Pc4;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88228080; PubMed=3372536;
RA   Ballard D.W., Philbrick W.M., Bothwell A.L.M.;
RT   "Identification of a novel 9-kDa polypeptide from nuclear extracts.
RT   DNA binding properties, primary structure, and in vitro expression.";
RL   J. Biol. Chem. 263:8450-8457(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Czech II; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: General coactivator that functions cooperatively with
CC       TAFs and mediates functional interactions between upstream
CC       activators and the general transcriptional machinery. Binds
CC       single-stranded DNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: Proteolysis is required for p9 DNA-binding activity.
CC   -!- PTM: Activity of p15 is controlled by protein kinases that target
CC       the regulatory domain. Phosphorylation inactivates cofactor
CC       function (By similarity).
CC   -!- SIMILARITY: Belongs to the transcriptional coactivator PC4 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03750; AAA37317.1; -.
DR   EMBL; BC010967; AAH10967.1; -.
DR   PIR; A28084; A28084.
DR   HSSP; P53999; 1PCF.
DR   TRANSFAC; T02138; -.
DR   MGD; MGI:104811; Rpo2tc1.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation;
KW   Transcription; Transcription regulation.
FT   INIT_MET      0      0       By similarity.
FT   PROPEP        1     49
FT   CHAIN        50    126       Activated RNA polymerase II
FT                                 transcriptional coactivator p15.
FT   DOMAIN        3     18       Ser-rich.
FT   DOMAIN       22     52       Lys-rich.
FT   DOMAIN       50     57       Ser-rich.
SQ   SEQUENCE   126 AA;  14296 MW;  7425871692FFC0EC CRC64;

  Query Match             32.5%;  Score 172;  DB 1;  Length 126;
  Best Local Similarity   41.4%;  Pred. No. 2.5e-09;
  Matches   36;  Conservative   20;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
              :|    ||       | |: || :  || | | |:|| : ||:::||||:: : :|:  ||
Db         40 KPGETSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPG 98

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |||| | |:||  |:: |  ||:|:::
Db         99 RKGISLNMEQWSQLKEQISDIDDAVRK 125


RESULT 8
BAC33314
ID   BAC33314    PRELIMINARY;      PRT;   127 AA.
AC   BAC33314;
DT   14-APR-2004 (TrEMBLrel. 27, Created)
DT   14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT   14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE   16 days embryo head cDNA, RIKEN full-length enriched library,
DE   clone:C130053F21 product:RNA polymerase II transcriptional
DE   coactivator, full insert sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK048371; BAC33314.1; -.
SQ   SEQUENCE   127 AA;  14427 MW;  12D58716F40FEB1C CRC64;

  Query Match             32.5%;  Score 172;  DB 2;  Length 127;
  Best Local Similarity   41.4%;  Pred. No. 2.6e-09;
  Matches   36;  Conservative   20;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
              :|    ||       | |: || :  || | | |:|| : ||:::||||:: : :|:  ||
Db         41 KPGETSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPG 99

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |||| | |:||  |:: |  ||:|:::
Db        100 RKGISLNMEQWSQLKEQISDIDDAVRK 126


RESULT 9
BAC33642
ID   BAC33642    PRELIMINARY;      PRT;   127 AA.
AC   BAC33642;
DT   14-APR-2004 (TrEMBLrel. 27, Created)
DT   14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT   14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE   ES cells cDNA, RIKEN full-length enriched library, clone:C330017N20
DE   product:RNA polymerase II transcriptional coactivator, full insert
DE   sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
```

```
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK049260; BAC33642.1; -.
SQ   SEQUENCE   127 AA;  14427 MW;  12D58716F40FEB1C CRC64;

  Query Match             32.5%;  Score 172;  DB 2;  Length 127;
  Best Local Similarity   41.4%;  Pred. No. 2.6e-09;
  Matches   36;  Conservative   20;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
              :|    ||       | |: || :  || | | |:|| : ||:::||||:: : :|:  ||
Db         41 KPGETSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPG 99

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |||| | |:||  |:: |  ||:|:::
Db        100 RKGISLNMEQWSQLKEQISDIDDAVRK 126


RESULT 10
TCP4_RAT
ID   TCP4_RAT       STANDARD;      PRT;   119 AA.
AC   Q63396;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Activated RNA polymerase II transcriptional coactivator p15 (Positive
DE   cofactor 4) (PC4) (p14) (Fragment).
GN   Name=Rpo2tc1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85046489; PubMed=6208900;
RA   Soma G.-I., Kitahara N., Andoh T.;
RT   "Molecular cloning and characterization of a cDNA clone for a protein
RT   specifically expressed in embryo as well as in a chemically induced
RT   pancreatic B cell tumor of rat.";
RL   Biochem. Biophys. Res. Commun. 124:164-171(1984).
CC   -!- FUNCTION: General coactivator that functions cooperatively with
CC       TAFs and mediates functional interactions between upstream
CC       activators and the general transcriptional machinery. Binds
CC       single-stranded DNA (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: Activity of p15 is controlled by protein kinases that target
CC       the regulatory domain. Phosphorylation inactivates cofactor
CC       function (By similarity).
CC   -!- SIMILARITY: Belongs to the transcriptional coactivator PC4 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02816; AAA41758.1; -.
DR   PIR; A23063; A23063.
DR   HSSP; P53999; 1PCF.
DR   TRANSFAC; T02139; -.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation;
KW   Transcription; Transcription regulation.
FT   NON_TER       1      1
FT   DOMAIN       <1     53       Regulatory.
FT   DOMAIN       <1     11       Ser-rich.
FT   DOMAIN       15     45       Lys-rich.
FT   DOMAIN       35     50       Ser-rich.
SQ   SEQUENCE   119 AA;  13657 MW;  D1D37FE79147635C CRC64;

  Query Match             31.3%;  Score 166;  DB 1;  Length 119;
  Best Local Similarity   40.2%;  Pred. No. 9.5e-09;
  Matches   35;  Conservative   21;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
              :|:   ||       | |: || :  || | | |:|| : ||:::||||:: : :|:  |
Db         33 KPSESSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPR 91

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
              |||| | |:||  |:: |  ||:|:::
Db         92 RKGISLNMEQWSQLKEQISDIDDAVRK 118


RESULT 11
Q9FY90
ID   Q9FY90      PRELIMINARY;      PRT;   110 AA.
AC   Q9FY90;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
```

```
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative transcriptional co-activator (At5g09240).
GN   Name=T5E8_40; Synonyms=At5g09240/T5E8_40;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA   Rudd S., Lemcke K., Mayer K.F.X.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA   Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA   Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA   Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA   Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL391712; CAC05450.1; -.
DR   EMBL; AK117904; BAC42542.1; -.
DR   EMBL; BT004696; AAO42942.1; -.
DR   HSSP; P53999; 1PCF.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0003713; F:transcription coactivator activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
SQ   SEQUENCE   110 AA;  12699 MW;  C835C3BB613F949D CRC64;

  Query Match             31.1%;  Score 165;  DB 2;  Length 110;
  Best Local Similarity   40.4%;  Pred. No. 1.1e-08;
  Matches   38;  Conservative   19;  Mismatches   23;  Indels   14;  Gaps    3;

Qy         20 RAAGDDGPSES----------AD--DDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDG 67
              ||: |   ||:          ||  :|| :  : ||||| ||  ||:: : ||:|: |||
Db         13 RASDDRDESETHAPPKKVAKPADEIEDIFICNLDKNRRVFVRNCNGRIWIAIRQFFVKDG 72

Qy         68 KTLP--GRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                |||   ::|| | ::||  ||:: : ||:|: |
Db         73 ITLPCNSKQGISLSLEQWNDLRNHEEDIDKALSE 106


RESULT 12
Q6E433
ID   Q6E433      PRELIMINARY;      PRT;    73 AA.
AC   Q6E433;
DT   01-OCT-2004 (TrEMBLrel. 28, Created)
DT   01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Activated RNA polymerase II transcription cofactor 4 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Simoes-Barbosa A., Teixeira A.R.L.;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY584192; AAT72821.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   73 AA;  8655 MW;  E566BF30D7B781CD CRC64;

  Query Match             30.5%;  Score 161.5;  DB 2;  Length 73;
  Best Local Similarity   45.1%;  Pred. No. 1.6e-08;
  Matches   32;  Conservative   19;  Mismatches   17;  Indels    3;  Gaps    2;

Qy         30 SADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPGRKGIQLPMDQWKILRD 88
                | ||::   || | | |:|| : ||:::||||:: : :|:  |||||| | |:||  |::
Db          4 SRDDNMF--QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPGRKGISLNMEQWSQLKE 61

Qy         89 NIKAIDEAIKE 99
                |  ||:|:::
Db         62 QISDIDDAVRK 72


RESULT 13
TCP4_HUMAN
ID   TCP4_HUMAN     STANDARD;      PRT;   126 AA.
AC   P53999; Q96L29;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Activated RNA polymerase II transcriptional coactivator p15 (Positive
DE   cofactor 4) (PC4) (p14).
GN   Name=RPO2TC1; Synonyms=PC4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94340741; PubMed=8062392;
RA   Kretzschmar M., Kaiser K., Lottspeich F., Meisterernst M.;
RT   "A novel mediator of class II gene transcription with homology to
RT   viral immediate-early transcriptional regulators.";
RL   Cell 78:525-534(1994).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-10 AND 80-96.
RX   MEDLINE=94340740; PubMed=8062391;
RA   Ge H., Roeder R.G.;
RT   "Purification, cloning, and characterization of a human coactivator,
RT   PC4, that mediates transcriptional activation of class II genes.";
RL   Cell 78:513-523(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow, Cervix, Lung, and Prostate;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   PHOSPHORYLATION, AND MASS SPECTROMETRY.
RX   MEDLINE=95108024; PubMed=7809103;
```

```
RA   Ge H., Zhao Y., Chait B.T., Roeder R.G.;
RT   "Phosphorylation negatively regulates the function of coactivator
RT   PC4.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:12691-12695(1994).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.74 ANGSTROMS) OF 62-126.
RX   MEDLINE=98025055; PubMed=9360603;
RA   Brandsen J., Werten S., van der Vliet P.C., Meisterernst M., Kroon J.,
RA   Gros P.;
RT   "C-terminal domain of transcription cofactor PC4 reveals dimeric ssDNA
RT   binding site.";
RL   Nat. Struct. Biol. 4:900-903(1997).
CC   -!- FUNCTION: General coactivator that functions cooperatively with
CC       TAFs and mediates functional interactions between upstream
CC       activators and the general transcriptional machinery. Binds
CC       single-stranded DNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: Activity of p15 is controlled by protein kinases that target
CC       the regulatory domain. Phosphorylation inactivates cofactor
CC       function. Seems to be phosphorylated in vivo by CK2 in at least 7
CC       sites in the N-terminal Ser-rich region.
CC   -!- MASS SPECTROMETRY: MW=14266; MW_ERR=4; METHOD=MALDI; RANGE=1-126;
CC       NOTE=Ref.4.
CC   -!- SIMILARITY: Belongs to the transcriptional coactivator PC4 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X79805; CAA56200.1; -.
DR   EMBL; U12979; AAA20980.1; -.
DR   EMBL; BC009610; AAH09610.1; -.
DR   EMBL; BC010537; AAH10537.1; -.
DR   EMBL; BC018189; AAH18189.1; -.
DR   EMBL; BC022339; AAH22339.1; -.
DR   PIR; A54670; A54670.
DR   PDB; 1PCF; X-ray; A/B/C/D/E/F/G/H=61-126.
DR   TRANSFAC; T02137; -.
DR   MIM; 600503; -.
DR   GO; GO:0005667; C:transcription factor complex; IDA.
DR   GO; GO:0003697; F:single-stranded DNA binding; IDA.
DR   GO; GO:0003713; F:transcription coactivator activity; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR003173; PC4.
DR   InterPro; IPR009044; ssDNA_bind_regul.
DR   Pfam; PF02229; PC4; 1.
KW   3D-structure; Activator; Direct protein sequencing; DNA-binding;
KW   Nuclear protein; Phosphorylation; Transcription;
KW   Transcription regulation.
FT   INIT_MET      0      0
FT   DOMAIN        1     60       Regulatory.
FT   DOMAIN        3     18       Ser-rich.
FT   DOMAIN       22     52       Lys-rich.
FT   DOMAIN       49     57       Ser-rich.
FT   CONFLICT     10     10       S -> G (in Ref. 3; AAH09610).
FT   STRAND       62     66
FT   TURN         67     68
FT   STRAND       69     76
FT   TURN         77     78
FT   STRAND       79     89
FT   TURN         91     92
FT   STRAND       95    104
FT   HELIX       106    126
SQ   SEQUENCE   126 AA;  14264 MW;  6B5E0CAACBAE65E5 CRC64;

  Query Match            30.3%;  Score 160.5;  DB 1;  Length 126;
  Best Local Similarity  37.9%;  Pred. No. 3.6e-08;
  Matches   36;  Conservative   21;  Mismatches   27;  Indels   11;  Gaps    3;

Qy         14 EPAAKRRAAGDDG--------PSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-E 64
              |    |::  |:            | | ||::   || | | |:|| : |||::||||:: :
Db         33 EKPVKKQKTGETSRALSSSKQSSSSRDDNMF--QIGKMRYVSVRDFKGKVLIDIREYWMD 90

Qy         65 KDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                :|:  |||||| |  :||  |:: |  ||:|:::
Db         91 PEGEMKPGRKGISLNPEQWSQLKEQISDIDDAVRK 125


RESULT 14
CAG33183
ID   CAG33183    PRELIMINARY;      PRT;   127 AA.
AC   CAG33183;
DT   01-JUN-2004 (TrEMBLrel. 27, Created)
DT   01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE   PC4 protein.
GN   PC4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RT   "Cloning of human full open reading frames in Gateway(TM) system entry
RT   vector (pDONR201).";
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR456902; CAG33183.1; -.
SQ   SEQUENCE   127 AA;  14395 MW;  0DAE0CAAAD5E4E15 CRC64;

  Query Match            30.3%;  Score 160.5;  DB 2;  Length 127;
  Best Local Similarity  37.9%;  Pred. No. 3.6e-08;
  Matches   36;  Conservative   21;  Mismatches   27;  Indels   11;  Gaps    3;

Qy         14 EPAAKRRAAGDDG--------PSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-E 64
              |    |::  |:            | | ||::   || | | |:|| : |||::||||:: :
Db         34 EKPVKKQKTGETSRALSSSKQSSSSRDDNMF--QIGKMRYVSVRDFKGKVLIDIREYWMD 91

Qy         65 KDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                :|:  |||||| |  :||  |:: |  ||:|:::
Db         92 PEGEMKPGRKGISLNPEQWSQLKEQISDIDDAVRK 126


RESULT 15
Q6DGP6
ID   Q6DGP6      PRELIMINARY;      PRT;   124 AA.
AC   Q6DGP6;
DT   01-OCT-2004 (TrEMBLrel. 28, Created)
DT   01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
```

```
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC076294; AAH76294.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   124 AA;  13888 MW;  DD624C883BC6EE52 CRC64;

  Query Match            30.2%;  Score 160;  DB 2;  Length 124;
  Best Local Similarity  39.6%;  Pred. No. 4e-08;
  Matches   36;  Conservative   22;  Mismatches   25;  Indels    8;  Gaps    3;

Qy         14 EPAAKRRAAGDDGPS-----ESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDG 67
              :|| |:::     ||     | : |::   || | | |:|| : |||::||||:: :: |
Db         34 KPAKKQKSGESSKPSGSAKTEKSSDNMF--QIGKLRYVSVRDFKGKVLIDIREYWMDQAG 91

Qy         68 KTLPGRKGIQLPMDQWKILRDNIKAIDEAIK 98
              :  ||:||| |  :||  |:: :  ||:|||
Db         92 EMKPGKKGISLNPEQWSQLKEQMSDIDDAIK 122


Search completed: January 10, 2005, 23:06:47
Job time : 196 secs
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 10, 2005, 22:58:22 ; Search time 38 Seconds
(without alignments)
176.266 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	117	22.1	292	4	US-09-538-092-608	Sequence 608, App
2	106	20.0	159	4	US-09-248-796A-19133	Sequence 19133, A
3	82.5	15.6	11877	3	US-09-105-537-6	Sequence 6, Appli
4	75	14.2	736	4	US-09-252-991A-31194	Sequence 31194, A
5	74	14.0	275	4	US-09-543-681A-7448	Sequence 7448, Ap
6	72	13.6	237	4	US-09-252-991A-17245	Sequence 17245, A
7	70.5	13.3	159	3	US-08-990-791-7	Sequence 7, Appli
8	70.5	13.3	159	4	US-09-372-591-7	Sequence 7, Appli
9	70	13.2	1404	3	US-08-801-308-1	Sequence 1, Appli
10	68.5	12.9	434	4	US-09-594-193-7	Sequence 7, Appli
11	68	12.8	525	2	US-08-676-166A-3	Sequence 3, Appli
12	67	12.6	499	4	US-09-252-991A-22825	Sequence 22825, A
13	66.5	12.5	416	4	US-09-270-767-43018	Sequence 43018, A
14	66	12.5	154	4	US-09-548-938A-19	Sequence 19, Appl
15	66	12.5	172	4	US-09-270-767-32192	Sequence 32192, A
16	66	12.5	172	4	US-09-270-767-47409	Sequence 47409, A
17	66	12.5	526	4	US-09-548-938A-10	Sequence 10, Appl
18	66	12.5	1016	4	US-09-625-972-24	Sequence 24, Appl
19	65.5	12.4	159	3	US-08-990-791-6	Sequence 6, Appli
20	65.5	12.4	159	4	US-09-372-591-6	Sequence 6, Appli
21	65.5	12.4	160	3	US-08-809-326A-14	Sequence 14, Appl
22	65.5	12.4	160	4	US-09-689-914A-14	Sequence 14, Appl
23	65.5	12.4	160	4	US-09-689-913A-14	Sequence 14, Appl
24	65.5	12.4	160	4	US-09-689-916A-14	Sequence 14, Appl
25	65.5	12.4	432	3	US-08-809-326A-16	Sequence 16, Appl
26	65.5	12.4	432	4	US-09-689-914A-16	Sequence 16, Appl
27	65.5	12.4	432	4	US-09-689-913A-16	Sequence 16, Appl
28	65.5	12.4	432	4	US-09-689-916A-16	Sequence 16, Appl
29	65.5	12.4	649	3	US-08-809-326A-15	Sequence 15, Appl
30	65.5	12.4	649	4	US-09-689-914A-15	Sequence 15, Appl
31	65.5	12.4	649	4	US-09-689-913A-15	Sequence 15, Appl
32	65.5	12.4	649	4	US-09-689-916A-15	Sequence 15, Appl
33	65.5	12.4	1967	4	US-09-849-602-16	Sequence 16, Appl
34	65	12.3	329	4	US-09-107-532A-3881	Sequence 3881, Ap
35	65	12.3	624	4	US-09-543-681A-4343	Sequence 4343, Ap
36	65	12.3	630	4	US-09-328-352-4238	Sequence 4238, Ap
37	65	12.3	1467	4	US-09-252-991A-17657	Sequence 17657, A
38	64.5	12.2	593	4	US-09-538-092-919	Sequence 919, App
39	64	12.1	216	1	US-07-917-722-4	Sequence 4, Appli
40	64	12.1	216	2	US-08-489-666C-5	Sequence 5, Appli
41	64	12.1	216	2	US-08-485-001B-5	Sequence 5, Appli
42	64	12.1	216	3	US-09-057-963A-4	Sequence 4, Appli
43	64	12.1	347	1	US-07-954-840A-4	Sequence 4, Appli
44	64	12.1	347	1	US-07-954-840A-6	Sequence 6, Appli
45	64	12.1	347	1	US-07-954-840A-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-538-092-608
; Sequence 608, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 608
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YMR039C
US-09-538-092-608

Query Match 22.1%; Score 117; DB 4; Length 292;
Best Local Similarity 32.3%; Pred. No. 1.2e-05;
Matches 32; Conservative 16; Mismatches 47; Indels 4; Gaps 2;

```
Qy          2 WRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIRE 61
              :|   :  |||  :      |      || |   :  : ||:|| || :    ::||||
Db          7 YRNKRRSDNGGGNLSNSNNNGGMPSGLSASD--AIFDLGKNKRVTVRQFRNINLIDIRE 64

Qy         62 FY--EKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIK 98
              :|     |:  ||:||| |  | :  |  :   ||||::
Db         65 YYLDSSTGEMKPGKKGISLTEDLYDELLKHRLNIDEALR 103
```

RESULT 2
US-09-248-796A-19133
; Sequence 19133, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

```
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19133
;   LENGTH: 159
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-19133

  Query Match             20.0%;  Score 106;  DB 4;  Length 159;
  Best Local Similarity   34.1%;  Pred. No. 0.00011;
  Matches   29;  Conservative   18;  Mismatches   30;  Indels    8;  Gaps    3;

Qy         16 AAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPGRK 74
              | || |      |: ::::  ::  : |  :| || :|   :|||||| ::||   ||:|
Db         15 AFKRGA-----PASNSNEVEILLDLKK--QVTVRKYNNVNLVDIREFYTDRDGVKRPGKK 67

Qy         75 GIQLPMDQWKILRDNIKAIDEAIKE 99
              || |  | :  | :    |  |: :
Db         68 GISLTEDTYYKLLEATNKIQNALDD 92


RESULT 3
US-09-105-537-6
; Sequence 6, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT: Liu, H.
;  APPLICANT: Xue, Y.
;  APPLICANT: Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/105,537A
;  CURRENT FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 11877
;   TYPE: PRT
;   ORGANISM: Streptomyces venezuelae
US-09-105-537-6

  Query Match             15.6%;  Score 82.5;  DB 3;  Length 11877;
  Best Local Similarity   26.9%;  Pred. No. 14;
  Matches   32;  Conservative   17;  Mismatches   35;  Indels   35;  Gaps    6;

Qy         10 GGGGEPAAKRR----AAGDDGP-----SESADDDIVV--------AQISKNRRVAVRTWN 52
              ||  | ||: |    | ||||       |::|||:          :    |    :| :
Db       8641 GGRPELAARLRALAAALGDDGDDATDLDEASDDDLFSFIDKELGDSDFMANNEDKLRDYL 8700

Qy         53 GKVVVDI----REFYEKDGKT-----------LPGRKGIQLPMDQWKILRDNIKAIDE 95
                :|  ::    |   | :|:|           |||  |:  | | |:::  :  || |
Db       8701 KRVTAELQQNTRRLREIEGRTHEPVAIVGMACRLPG--GVASPEDLWQLVAGDGDAISE 8757


RESULT 4
US-09-252-991A-31194
; Sequence 31194, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 31194
;   LENGTH: 736
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31194

  Query Match             14.2%;  Score 75;  DB 4;  Length 736;
  Best Local Similarity   34.7%;  Pred. No. 3.2;
  Matches   26;  Conservative    9;  Mismatches   20;  Indels   20;  Gaps    4;

Qy          2 WRKGNKRFGG---GGE------PAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWN 52
              :| |  |||   |||         |:|||| |  :  | | :|  ::       ||
Db        636 FRGGRLRFGGSGAGGERDAEEGEGAERRAAGQDLAAAEAGDQVVHVRV-------VRMVA 688

Qy         53 GKVVVDIREFYEKDG 67
              |::|       |:||
Db        689 GQLVA----ITEQDG 699


RESULT 5
US-09-543-681A-7448
; Sequence 7448, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7448
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Proteus mirabilis
US-09-543-681A-7448

  Query Match             14.0%;  Score 74;  DB 4;  Length 275;
  Best Local Similarity   29.5%;  Pred. No. 1.2;
  Matches   18;  Conservative   14;  Mismatches   23;  Indels    6;  Gaps    2;

Qy         39 QISKNRRVAVRTW--NGKVVVDIREFYEKD----GKTLPGRKGIQLPMDQWKILRDNIKA 92
              | |||:::|:: |  : |:||  :   |  :       :| ||   |   :  :| : |||
Db        146 QQSKNKKIAIKPWLMDNKLVVGVGNIYANEALFSSGIMPDRKANSLTEQECDVLVNAIKA 205

Qy         93 I 93
              :
Db        206 V 206


RESULT 6
US-09-252-991A-17245
; Sequence 17245, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 17245
;   LENGTH: 237
;   TYPE: PRT
```

```
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17245

  Query Match             13.6%;  Score 72;  DB 4;  Length 237;
  Best Local Similarity   28.1%;  Pred. No. 1.7;
  Matches   25;  Conservative   11;  Mismatches   19;  Indels   34;  Gaps    5;

Qy          2 WRKG------NKRFG----GGGEPAAKRRAAGDDGPSESADDDIV----VAQISKNRRVA 47
              |  |         :| |    | |:|| :||       | | ||||::    : :    ||||
Db          6 WGPGEDDQYQQQRLGAQVAGDGDPAEQRRRR----PGEPADDDVLRRGALQEAGVERRVA 61

Qy         48 VRTWNGKVVVDIREFYEKDGKTLPGRKGI 76
                          |: |:  |||:' :
Db         62 ----------------EQRGEGQPGRQRV 74


RESULT 7
US-08-990-791-7
; Sequence 7, Application US/08990791
; Patent No. 6229001
; GENERAL INFORMATION:
;  APPLICANT: Southern Research Institute
;  TITLE OF INVENTION: MYCOBACTERIUM FOR A GENE THAT ENCODES FOR THE ENZYME
;  TITLE OF INVENTION: DIHYDROFOLATE REDUCTASE
;  FILE REFERENCE: 063560.0102
;  CURRENT APPLICATION NUMBER: US/08/990,791
;  CURRENT FILING DATE: 1997-12-15
;  EARLIER APPLICATION NUMBER: 60/034,725
;  EARLIER FILING DATE: 1997-01-03
;  EARLIER APPLICATION NUMBER: 60/039,737
;  EARLIER FILING DATE: 1997-02-14
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 7
;   LENGTH: 159
;   TYPE: PRT
;   ORGANISM: Citrobacter freundii
US-08-990-791-7

  Query Match             13.3%;  Score 70.5;  DB 3;  Length 159;
  Best Local Similarity   27.1%;  Pred. No. 1.5;
  Matches   23;  Conservative   13;  Mismatches   20;  Indels   29;  Gaps    4;

Qy         36 VVAQISKNRRVAVRT---WN-------------GKVVVDIREFYEKDGKTLPGRKGIQLP 79
              ::| :: :| : :     ||             | ||  |  :|  |: ||||| | :
Db          4 LIAALAVDRVIGMENAMPWNLPADLAWFKRNTLNKPVVMGRHTWESIGRPLPGRKNIVIS 63

Qy         80 MD-------QWKILRDNIKAIDEAI 97
                    ||      :|::||||
Db         64 SKPGTDDRVQW------VKSVDEAI 82


RESULT 8
US-09-372-591-7
; Sequence 7, Application US/09372591
; Patent No. 6559295
; GENERAL INFORMATION:
;  APPLICANT: Southern Research Institute
;  TITLE OF INVENTION: MYCOBACTERIUM FOL A GENE THAT ENCODES FOR THE ENZYME
;  TITLE OF INVENTION: DIHYDROFOLATE REDUCTASE
;  FILE REFERENCE: 063560.0108
;  CURRENT APPLICATION NUMBER: US/09/372,591
;  CURRENT FILING DATE: 1999-08-12
;  EARLIER APPLICATION NUMBER: 60/034,725
;  EARLIER FILING DATE: 1997-01-03
;  EARLIER APPLICATION NUMBER: 60/039,737
;  EARLIER FILING DATE: 1997-02-14
;  EARLIER APPLICATION NUMBER: 08/990,791
;  EARLIER FILING DATE: 1997-12-15
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 159
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer that can hybridise to flanking regions of
;   OTHER INFORMATION: pig microsatellites sequences
;   FEATURE:
;   OTHER INFORMATION: Primer that can hybridise to flanking regions of
;   OTHER INFORMATION: pig microsatellites sequences
;   FEATURE:
;   OTHER INFORMATION: Primer that can hybridise to flanking regions of
;   OTHER INFORMATION: pig microsatellites sequences
US-09-372-591-7

  Query Match             13.3%;  Score 70.5;  DB 4;  Length 159;
  Best Local Similarity   27.1%;  Pred. No. 1.5;
  Matches   23;  Conservative   13;  Mismatches   20;  Indels   29;  Gaps    4;

Qy         36 VVAQISKNRRVAVRT---WN-------------GKVVVDIREFYEKDGKTLPGRKGIQLP 79
              ::| :: :| : :     ||             | ||  |  :|  |: ||||| | :
Db          4 LIAALAVDRVIGMENAMPWNLPADLAWFKRNTLNKPVVMGRHTWESIGRPLPGRKNIVIS 63

Qy         80 MD-------QWKILRDNIKAIDEAI 97
                    ||      :|::||||
Db         64 SKPGTDDRVQW------VKSVDEAI 82


RESULT 9
US-08-801-308-1
; Sequence 1, Application US/08801308
; Patent No. 6368790
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, Robert E.
;    TITLE OF INVENTION:  cDNA ENCODING P2P PROTEINS AND USE OF
;    TITLE OF INVENTION:  P2P cDNA-DERIVED ANTIBODIES AND ANTISENSE REAGENTS IN
;    TITLE OF INVENTION:  DETERMINING THE PROLIFERATIVE POTENTIAL OF NORMAL,
;    TITLE OF INVENTION:  ABNORMAL AND CANCER CELLS IN ANIMALS AND HUMANS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates, P.C.
;      STREET:  230 S. Fifteenth Street, Suite 500
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/801,308
;      FILING DATE:  18-FEB-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiser, Gerard J.
;      REGISTRATION NUMBER:  19,763
;      REFERENCE/DOCKET NUMBER:  372.6435P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-875-8383
;      TELEFAX:  215-875-8394
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1404 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-801-308-1

  Query Match             13.2%;  Score 70;  DB 3;  Length 1404;
  Best Local Similarity   27.6%;  Pred. No. 28;
  Matches   24;  Conservative   16;  Mismatches   39;  Indels    8;  Gaps    4;

Qy         17 AKRRAAGDDGPSESADDDIVVAQI--SKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRK 74
```

```
              |||: || :| | :  |    :    |  |:   :|   : |:   | |  |  | |
Db       1039 AKRKVAGSEGSSSTLVDYTSTSSTGGSPVRKSEEKTDTKRTVIKTMEEYNND-NTAPAED 1097

Qy         75 ---GIQLPMDQWKILRDNIKAIDEAIK 98
                 ||:|  :|   :|: :: :| :|
Db       1098 VIIMIQVPQSKWD--KDDFESEEEDVK 1122


RESULT 10
US-09-594-193-7
; Sequence 7, Application US/09594193
; Patent No. 6441274
; GENERAL INFORMATION:
;  APPLICANT:  CAHOON, REBECCA E.
;  APPLICANT:  FALCO, S. CARL
;  APPLICANT:  LOHMAN, KARIN
;  TITLE OF INVENTION:  NOVEL PLANT TRYPTOPHAN SYNTHASE BETA SUBUNIT
;  FILE REFERENCE:  BB1374 US NA
;  CURRENT APPLICATION NUMBER: US/09/594,193
;  CURRENT FILING DATE:  2000-06-15
;  PRIOR APPLICATION NUMBER:  60/139,568
;  PRIOR FILING DATE:  1999-06-16
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 434
;   TYPE: PRT
;   ORGANISM: Aquifex aeolicus
US-09-594-193-7

  Query Match             12.9%;  Score 68.5;  DB 4;  Length 434;
  Best Local Similarity   28.2%;  Pred. No. 9.4;
  Matches   22;  Conservative   14;  Mismatches   15;  Indels   27;  Gaps    4;

Qy         35 IVVAQISKN----RRVAVRTWNGKVVVDI-------REFYEKDGKTLPGRKGIQLPMDQW 83
              : : ::| |     ||: : || |:|:          |::||:: :  ||  ||
Db        155 VYMVRVSYNQKPYRRILMETWKGEVIPSPSPYTNAGRKYYEENPEH-PGSLGI------- 206

Qy         84 KILRDNIKAIDEAIKENA 101
                  || |||:| |
Db        207 --------AISEAIEEAA 216


RESULT 11
US-08-676-166A-3
; Sequence 3, Application US/08676166A
; Patent No. 5955270
;  GENERAL INFORMATION:
;    APPLICANT:  Radford, Alan
;    APPLICANT:  Parish, John H.
;    TITLE OF INVENTION:  EXPLOITATION OF THE CELLULASE COMPLEX OF
;    TITLE OF INVENTION:  NEUROSPORA
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David A. Jackson, Esq.
;      STREET:  411 Hackensack Ave, Continental Plaza, 4th
;      STREET:  Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/676,166A
;      FILING DATE:  15-JUL-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  1321-1-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  525 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  H. grisea
US-08-676-166A-3

  Query Match             12.8%;  Score 68;  DB 2;  Length 525;
  Best Local Similarity   28.1%;  Pred. No. 14;
  Matches   25;  Conservative   12;  Mismatches   24;  Indels   28;  Gaps    5;

Qy          2 WRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIRE 61
              :|:||| | | |             ::      || |  |:          ||  : :|:
Db        285 YRQGNKTFYGKGMTV-----------DTTKKITVVTQFLKDA-------NGD-LGEIKR 324

Qy         62 FYEKDGK-------TLPGRKGIQLPMDQW 83
              || :|||       |:|| :|  :  | |
Db        325 FYVQDGKIIPNSESTIPGVEGNSITQD-W 352


RESULT 12
US-09-252-991A-22825
; Sequence 22825, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 22825
;   LENGTH: 499
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22825

  Query Match             12.6%;  Score 67;  DB 4;  Length 499;
  Best Local Similarity   29.2%;  Pred. No. 17;
  Matches   26;  Conservative   14;  Mismatches   29;  Indels   20;  Gaps    5;

Qy          4 KGNKRFGGGGEPAAKRRAAGDD-----GPSESADDDIVVA-----QISKNRRVAVRTWNG 53
              | |    | ||||::  | |::     ||||   | | :|      : |:  |
Db        170 KDNAEEEGEGEPASEEVAEGEEPTRIPGPSEK--DGIKIAIIGRPNVGKSTLVNRMLGEE 227

Qy         54 KVVVDIREFYEKDGKTLPGRKGIQLPMDQ 82
              :|:|     |:: | |   |  | :| ::
Db        228 RVIV-----YDQAGTT---RDSIYIPFER 248


RESULT 13
US-09-270-767-43018
; Sequence 43018, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
```

```
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43018
;   LENGTH: 416
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-43018

  Query Match             12.5%;  Score 66.5;  DB 4;  Length 416;
  Best Local Similarity   22.9%;  Pred. No. 15;
  Matches   33;  Conservative   13;  Mismatches   51;  Indels   47;  Gaps    7;

Qy          3 RKGNKRFGG--GGE-PAAKRRAAGDDGP--------------SESADDDIVVAQISKNR 44
              | |||   |  | | |  :    : ||                :|  |   |  | |
Db        111 RYGNKESNGEFGNELPQRQFNNRDNRGPPRVRTGEKFGKREFDRQSGSDRTGVKSIDKRE 170

Qy         45 RVAVRTWNGKVVVDIREFYEKDGKTLPGR--------------------KGIQLPMDQWK 84
                 | |    || :  :  |:| |                        :  |: :|:||
Db        171 GGGAHNW-GSPKQDIEDL-KTTGETSPQAEKEDSANEQSADPAVAAEEDESKQMTLDEWK 228

Qy         85 ILRD-------NIKAIDEAIKENA 101
               |||       |::   |   :||
Db        229 ALRDQRAKPNYNLRKAGEGAADNA 252


RESULT 14
US-09-548-938A-19
; Sequence 19, Application US/09548938A
; Patent No. 6573086
; GENERAL INFORMATION:
;  APPLICANT: EMALFARB, MARK AARON
;  APPLICANT:  BURLINGAME, RICHARD PAUL
;  APPLICANT:  OLSON, PHILIP TERRY
;  APPLICANT:  SINITSYN, ARKADY PANTELEIMONOVICK
;  APPLICANT:  PARRICHE, MARTINE
;  APPLICANT:  BOUSSON, JEAN CHRISTOPHE
;  APPLICANT:  PYNNONEN, CHRISTINE MARIE
;  APPLICANT:  PUNT, PETER JAN
;  APPLICANT:  VAN-ZEIJL, CORNELIA MARIA JOHANNA
;  TITLE OF INVENTION: TRANSFORMATION SYSTEM IN THE FIELD OF FILAMENTOUS FUNGI
;  FILE REFERENCE: 3123-4001
;  CURRENT APPLICATION NUMBER: US/09/548,938A
;  CURRENT FILING DATE:  2000-04-13
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Chrysosporium sp.
US-09-548-938A-19

  Query Match             12.5%;  Score 66;  DB 4;  Length 154;
  Best Local Similarity   25.8%;  Pred. No. 5;
  Matches   23;  Conservative   13;  Mismatches   25;  Indels   28;  Gaps    4;

Qy          2 WRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIRE 61
              :|:||| | | |                ::     || |  ||       :   : :|:
Db         44 YRQGNKTFYGKGMTV------------DTTKKITVVTQFLKN--------SAGELSEIKR 83

Qy         62 FYEKDGK-------TLPGRKGIQLPMDQW 83
              || ::||       |:|| :|  :  | |
Db         84 FYVQNGKVIPNSESTIPGVEGNSITQD-W 111


RESULT 15
US-09-270-767-32192
; Sequence 32192, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32192
;   LENGTH: 172
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-32192

  Query Match             12.5%;  Score 66;  DB 4;  Length 172;
  Best Local Similarity   27.2%;  Pred. No. 5.7;
  Matches   22;  Conservative   10;  Mismatches   39;  Indels   10;  Gaps    3;

Qy          2 WRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIRE 61
              |::     |    |   |   | ||   : |       : |   ::|:|    |||: :
Db         53 WKRSLSHVGSA--PDLSRVIVGKDGFEANVD-----VHLFKPYEISVKTSGDTVVVEAKH 105

Qy         62 FYEKDGKTLPGR---KGIQLP 79
                :|| |  ||   |   ||
Db        106 EKRRDGDTFVGRHIVKRFVLP 126


Search completed: January 10, 2005, 23:08:15
Job time : 39 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 10, 2005, 23:06:53 ; Search time 144 Seconds
(without alignments)
252.901 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1603904 seqs, 360571292 residues

Total number of hits satisfying chosen parameters: 1603904

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	530	100.0	475	16	US-10-437-963-185855	Sequence 185855,
2	398	75.1	119	15	US-10-425-114-42612	Sequence 42612, A
3	395	74.5	103	17	US-10-425-115-259005	Sequence 259005,
4	269	50.8	105	15	US-10-424-599-181609	Sequence 181609,
5	261.5	49.3	105	15	US-10-424-599-181610	Sequence 181610,
6	182.5	34.4	163	15	US-10-424-599-147198	Sequence 147198,
7	182.5	34.4	196	15	US-10-425-114-38036	Sequence 38036, A
8	180.5	34.1	184	16	US-10-437-963-162084	Sequence 162084,
9	178.5	33.7	367	15	US-10-424-599-151999	Sequence 151999,
10	177	33.4	79	15	US-10-424-599-270637	Sequence 270637,
11	172	32.5	141	13	US-10-087-192-1785	Sequence 1785, Ap
12	165	31.1	140	17	US-10-425-115-205102	Sequence 205102,
13	165	31.1	150	15	US-10-425-114-67613	Sequence 67613, A
14	165	31.1	183	17	US-10-425-115-339084	Sequence 339084,
15	165	31.1	198	15	US-10-425-114-60171	Sequence 60171, A
16	165	31.1	204	15	US-10-425-114-69791	Sequence 69791, A
17	160.5	30.3	127	13	US-10-087-192-1788	Sequence 1788, Ap
18	160.5	30.3	127	14	US-10-177-293-350	Sequence 350, App
19	160.5	30.3	127	14	US-10-257-021-22	Sequence 22, Appl
20	160.5	30.3	127	17	US-10-370-715B-282	Sequence 282, App
21	160.5	30.3	132	9	US-09-925-297-779	Sequence 779, App
22	153.5	29.0	106	16	US-10-767-701-60649	Sequence 60649, A
23	110	20.8	104	11	US-09-864-408A-4696	Sequence 4696, Ap
24	107.5	20.3	82	17	US-10-425-115-342501	Sequence 342501,
25	101.5	19.2	88	17	US-10-425-115-251977	Sequence 251977,
26	87	16.4	124	17	US-10-425-115-202622	Sequence 202622,
27	83.5	15.8	87	17	US-10-425-115-272770	Sequence 272770,
28	82.5	15.6	11877	9	US-09-861-289-6	Sequence 6, Appli
29	82.5	15.6	11877	9	US-09-860-846-6	Sequence 6, Appli
30	82.5	15.6	11877	10	US-09-836-821-6	Sequence 6, Appli
31	82.5	15.6	11877	14	US-10-271-889-49	Sequence 49, Appl
32	82.5	15.6	12199	10	US-09-988-384B-6	Sequence 6, Appli
33	80	15.1	118	17	US-10-425-115-334183	Sequence 334183,
34	79.5	15.0	115	16	US-10-767-701-34940	Sequence 34940, A
35	79.5	15.0	134	17	US-10-425-115-286086	Sequence 286086,
36	79.5	15.0	146	17	US-10-425-115-341160	Sequence 341160,
37	76.5	14.4	189	16	US-10-437-963-104003	Sequence 104003,
38	76	14.3	507	16	US-10-437-963-127962	Sequence 127962,
39	76	14.3	924	17	US-10-425-115-358126	Sequence 358126,
40	75.5	14.2	499	15	US-10-282-122A-61693	Sequence 61693, A
41	74.5	14.1	100	16	US-10-767-701-37709	Sequence 37709, A
42	74.5	14.1	465	15	US-10-282-122A-47419	Sequence 47419, A
43	74	14.0	506	16	US-10-437-963-164624	Sequence 164624,
44	73.5	13.9	128	17	US-10-425-115-336013	Sequence 336013,
45	73	13.8	263	16	US-10-325-468-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-10-437-963-185855
; Sequence 185855, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 185855
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8270C.1.pep
US-10-437-963-185855

Query Match 100.0%; Score 530; DB 16; Length 475;
Best Local Similarity 100.0%; Pred. No. 8.5e-52;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIR 434

Qy         61 EFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
```

```
                ||||||||||||||||||||||||||||||||||||||||
Db        435 EPYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 475
```

RESULT 2
US-10-425-114-42612
; Sequence 42612, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42612
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700380835_FLI.pep
US-10-425-114-42612

Query Match 75.1%; Score 398; DB 15; Length 119;
Best Local Similarity 74.5%; Pred. No. 1.9e-37;
Matches 76; Conservative 13; Mismatches 11; Indels 2; Gaps 1;

```
Qy          1 MWRKGNKRPGGG--GEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVD 58
              || || ||||||  |||||||:|| ||||||||:|  |||:||||::|:||:| |:| ||
Db         17 MWGKGKKRPGGGGSGEPAAKRQAARDDGPSESAEDGTVVAEISKNKKVSVRSWKGRVFVD 76

Qy         59 IREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKEN 100
              :|||| ||||||| |||| | :||||||:||||||:|||:||
Db         77 LREFYFKDGKTLPTRKGISLQLDQWKILKDNIKAINEAIEEN 118
```

RESULT 3
US-10-425-115-259005
; Sequence 259005, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 259005
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167799C.1.pep
US-10-425-115-259005

Query Match 74.5%; Score 395; DB 17; Length 103;
Best Local Similarity 73.5%; Pred. No. 3.4e-37;
Matches 75; Conservative 14; Mismatches 11; Indels 2; Gaps 1;

```
Qy          1 MWRKGNKRPGGG--GEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVD 58
              || :| ||||||  |||||||:|| ||||||||:|  |||:||||::|:||:| |:| ||
Db          1 MWGRGKKRPGGGGSGEPAAKRQAARDDGPSESAEDGTVVAEISKNKKVSVRSWKGRVFVD 60

Qy         59 IREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKEN 100
              :|||| ||||||| |||| | :||||||:||||||:|||:||
Db         61 LREFYFKDGKTLPTRKGISLQLDQWKILKDNIKAINEAIEEN 102
```

RESULT 4
US-10-424-599-181609
; Sequence 181609, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181609
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135006C.1.pep
US-10-424-599-181609

Query Match 50.8%; Score 269; DB 15; Length 105;
Best Local Similarity 54.4%; Pred. No. 9.2e-23;
Matches 56; Conservative 13; Mismatches 14; Indels 20; Gaps 3;

```
Qy         17 AKRRAAGDDGPSES--------------ADDD---IVVAQISKNRRVAVRTWNGKVVVD 58
              ||||   ||| |::                :|||   | | :||||||||||| | | ::||
Db          5 AKRR--DDDGASDADSEGHAPPKKSLKKDSDDDPDSITVCEISKNRRVAVRNWKGSIMVD 62

Qy         59 IREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
              ||||| |||| ||||||| | |||| :||:::: ||:|| ||:
Db         63 IREFYVKDGKQLPGRKGISLTMDQWNVLRNHVEEIDKAINENS 105
```

RESULT 5
US-10-424-599-181610
; Sequence 181610, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181610
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135007C.1.pep
US-10-424-599-181610

Query Match 49.3%; Score 261.5; DB 15; Length 105;
Best Local Similarity 52.7%; Pred. No. 6.6e-22;
Matches 48; Conservative 14; Mismatches 24; Indels 5; Gaps 1;

```
Qy         11 GGGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTL 70
              |   |    :  || |     | : : :|||||||||| | | ::|||||| |||| |
Db         20 GHAPPMKSLKKDSDDDP-----DSVTICEISKNRRVAVRNWKGSIMVDIREFYVKDGKQL 74

Qy         71 PGRKGIQLPMDQWKILRDNIKAIDEAIKENA 101
```

```
                ||||||  |  ||||  :||:::: ||:|: ||:
Db         75 PGRKGISLTMDQWNVLRNHVEEIDKAVNENS 105


RESULT 6
US-10-424-599-147198
; Sequence 147198, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147198
;   LENGTH: 163
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_10393C.1.pep
US-10-424-599-147198

  Query Match             34.4%;  Score 182.5;  DB 15;  Length 163;
  Best Local Similarity   40.9%;  Pred. No. 1.3e-12;
  Matches   36;  Conservative   20;  Mismatches   23;  Indels    9;  Gaps    2;

Qy         12 GGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLP 71
              ||  |:| :   |:|       |::: ::|  ||| :: : || :| |||:|:|||| ||
Db         81 GG--ASKGKEYDDEG-------DLIICRLSDKRRVTIQDFRGKTLVSIREYYKKDGKELP 131

Qy         72 GRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                ||| |  :||   : |: ||::|||:
Db        132 TSKGISLTEEQWSAFKKNVPAIEKAIKK 159


RESULT 7
US-10-425-114-38036
; Sequence 38036, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38036
;   LENGTH: 196
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3242-699-H3_FLI.pep
US-10-425-114-38036

  Query Match             34.4%;  Score 182.5;  DB 15;  Length 196;
  Best Local Similarity   40.9%;  Pred. No. 1.6e-12;
  Matches   36;  Conservative   20;  Mismatches   23;  Indels    9;  Gaps    2;

Qy         12 GGEPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLP 71
              ||  |:| :   |:|       |::: ::|  ||| :: : || :| |||:|:|||| ||
Db        114 GG--ASKGKEYDDEG-------DLIICRLSDKRRVTIQDFRGKTLVSIREYYKKDGKELP 164

Qy         72 GRKGIQLPMDQWKILRDNIKAIDEAIKE 99
                ||| |  :||   : |: ||::|||:
Db        165 TSKGISLTEEQWSAFKKNVPAIEKAIKK 192


RESULT 8
US-10-437-963-162084
; Sequence 162084, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 162084
;   LENGTH: 184
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_61208C.1.pep
US-10-437-963-162084

  Query Match             34.1%;  Score 180.5;  DB 16;  Length 184;
  Best Local Similarity   36.5%;  Pred. No. 2.5e-12;
  Matches   38;  Conservative   22;  Mismatches   29;  Indels   15;  Gaps    3;

Qy         10 GGGGEPAAKRRAAGDD-----------GPSESADD--DIVVAQISKNRRVAVRTWNGKV 55
              || ||     :|  ::           |     ||  |::: ::|  ||| :: : ||
Db         76 GGAGEEGKDEQAEEEEEEEEEEEEERGGGTKREYDDEGDLILCRLSARRRVTLQEFKGKT 135

Qy         56 VVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
              :: |||:| |||| ||  ||| | ::||:  ||:: ||::|||:
Db        136 LLSIREYYFKDGKELPA-KGISLTVEQWEAFRDSVPAIEDAIKK 178


RESULT 9
US-10-424-599-151999
; Sequence 151999, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 151999
;   LENGTH: 367
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_108279C.1.pep
US-10-424-599-151999

  Query Match             33.7%;  Score 178.5;  DB 15;  Length 367;
  Best Local Similarity   36.1%;  Pred. No. 1e-11;
  Matches   39;  Conservative   19;  Mismatches   39;  Indels   11;  Gaps    2;

Qy          3 RKGNKRFGGGGEPAAKRR---AAGDDGPSESA--------DDDIVVAQISKNRRVAVRTW 51
```

```
Db         86 KKENEDIAAKNDDVAKKEDVVVANEEESRETEVLPKLKRDDPERVICHLSNRRNVAVKDF 145

Qy         52 NGKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKE 99
Db        146 KGTTLVSIREFYMKDGKPLPGSKGISLSSEQWSTFKKSVPAIEEAIKK 193


RESULT 10
US-10-424-599-270637
; Sequence 270637, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 270637
;   LENGTH: 79
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_86402C.1.pep
US-10-424-599-270637

  Query Match            33.4%;  Score 177;  DB 15;  Length 79;
  Best Local Similarity  41.9%;  Pred. No. 2.2e-12;
  Matches   31;  Conservative   18;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         26 GPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKI 85
Db          2 GKEYDDEGDLIICRLSDKRRVTIQDFRGKTLVSIREYYKKDGKELPTSKGISLTEEQWST 61

Qy         86 LRDNIKAIDEAIKE 99
Db         62 FKKNVPAIEKAIKK 75


RESULT 11
US-10-087-192-1785
; Sequence 1785, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1785
;   LENGTH: 141
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-087-192-1785

  Query Match            32.5%;  Score 172;  DB 13;  Length 141;
  Best Local Similarity  41.4%;  Pred. No. 1.7e-11;
  Matches   36;  Conservative   20;  Mismatches   29;  Indels    2;  Gaps    2;

Qy         14 EPAAKRRAAGDDGPSESADDDIVVAQISKNRRVAVRTWNGKVVVDIREFY-EKDGKTLPG 72
Db         55 KPGETSRALASSKQSSSSRDDNMF-QIGKMRYVSVRDFKGKILIDIREYWMDSEGEMKPG 113

Qy         73 RKGIQLPMDQWKILRDNIKAIDEAIKE 99
Db        114 RKGISLNMEQWSQLKEQISDIDDAVRK 140


RESULT 12
US-10-425-115-205102
; Sequence 205102, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 205102
;   LENGTH: 140
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_118642C.1.pep
US-10-425-115-205102

  Query Match            31.1%;  Score 165;  DB 17;  Length 140;
  Best Local Similarity  40.5%;  Pred. No. 1.1e-10;
  Matches   32;  Conservative   17;  Mismatches   28;  Indels    2;  Gaps    1;

Qy         23 GDDGPSESADD--DIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPM 80
Db         57 GKGGGKREYDDQGDLILCRLSSKRRVTLSEFKGRSLVSIREFYVKDGKEMPSAKGISMTM 116

Qy         81 DQWKILRDNIKAIDEAIKE 99
Db        117 EQWEAFCNAVPAIEAAIKK 135


RESULT 13
US-10-425-114-67613
; Sequence 67613, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67613
;   LENGTH: 150
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3279-041-H1_FLI.pep
US-10-425-114-67613

  Query Match            31.1%;  Score 165;  DB 15;  Length 150;
  Best Local Similarity  40.5%;  Pred. No. 1.1e-10;
  Matches   32;  Conservative   17;  Mismatches   28;  Indels    2;  Gaps    1;
```

```
Qy         23 GDDGPSESADD--DIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPM 80
              |  |      ||  |::: ::|  ||| :  : |: :| ||||| |||| :|  ||| : |
Db         67 GKGGGKREYDDQGDLILCRLSSKRRVTLSEFKGRSLVSIREFYVKDGKEMPSAKGISMTM 126

Qy         81 DQWKILRDNIKAIDEAIKE 99
              :||:   : : ||: |||:
Db        127 EQWEAFCNAVPAIEAAIKK 145


RESULT 14
US-10-425-115-339084
; Sequence 339084, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 339084
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_72412C.1.pep
US-10-425-115-339084

  Query Match             31.1%;  Score 165;  DB 17;  Length 183;
  Best Local Similarity   39.2%;  Pred. No. 1.5e-10;
  Matches   31;  Conservative   19;  Mismatches   27;  Indels    2;  Gaps    1;

Qy         23 GDDGPSESADD--DIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPM 80
              |  |      ||  |::: ::|  ||| :  : |: :| ||||| |||| :|  ||| : :
Db        100 GKGGGKREYDDQGDLILCRLSSKRRVTLSEFKGRSLVSIREFYVKDGKEMPSAKGISMTL 159

Qy         81 DQWKILRDNIKAIDEAIKE 99
              :||:   : : ||::|||:
Db        160 EQWEAFCNAVPAIEDAIKK 178


RESULT 15
US-10-425-114-60171
; Sequence 60171, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60171
;   LENGTH: 198
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3608-031-G8_FLI.pep
US-10-425-114-60171

  Query Match             31.1%;  Score 165;  DB 15;  Length 198;
  Best Local Similarity   39.2%;  Pred. No. 1.6e-10;
  Matches   31;  Conservative   19;  Mismatches   27;  Indels    2;  Gaps    1;

Qy         23 GDDGPSESADD--DIVVAQISKNRRVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPM 80
              |  |      ||  |::: ::|  ||| :  : |: :| ||||| |||| :|  ||| : :
Db        115 GKGGGKREYDDQGDLILCRLSSKRRVTLSEFKGRSLVSIREFYVKDGKEMPSAKGISMTL 174

Qy         81 DQWKILRDNIKAIDEAIKE 99
              :||:   : : ||::|||:
Db        175 EQWEAFCNAVPAIEDAIKK 193


Search completed: January 10, 2005, 23:19:10
Job time : 145 secs
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 11, 2005, 23:15:54 ; Search time 3302 Seconds
(without alignments)
1446.475 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10629953/runat_10012005_172535_9414/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10629953_@CGN_1_1_3731_@runat_10012005_172535_9414 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	530	100.0	691	8	AK072440	AK072440 Oryza sat
2	368	69.4	289	6	AR246100	AR246100 Sequence
3	352	66.4	145419	8	AP004876	AP004876 Oryza sat
4	297	56.0	263	6	AR246253	AR246253 Sequence
5	275	51.9	286	6	AR251557	AR251557 Sequence
6	247.5	46.7	324	8	BT004745	BT004745 Arabidops
7	247.5	46.7	605	8	AF053302	AF053302 Arabidops
8	185	34.9	913	8	AB050390	AB050390 Brassica
9	183	34.5	498	8	BT005847	BT005847 Arabidops
10	183	34.5	730	8	AF053303	AF053303 Arabidops
11	183	34.5	757	8	AY085896	AY085896 Arabidops
12	182	34.3	923	8	AK104192	AK104192 Oryza sat
13	182	34.3	945	8	AK070545	AK070545 Oryza sat
14	182	34.3	1141	8	AK059936	AK059936 Oryza sat
c 15	174	32.8	82896	8	ATT2K12	AL590346 Arabidops
c 16	174	32.8	103534	8	ATT5E8	AL391712 Arabidops
17	172	32.5	694	10	BC010967	BC010967 Mus muscu
18	172	32.5	2777	10	MUSBPP9	J03750 Mouse singl
19	169	31.9	333	8	BT004696	BT004696 Arabidops
20	169	31.9	538	8	AK117904	AK117904 Arabidops
21	166	31.3	615	10	RATONC	K02816 Rat unident
22	165	31.1	678	5	CR405748	CR405748 Gallus ga
23	165	31.1	685	5	BX930917	BX930917 Gallus ga
24	161.5	30.5	470	9	AY584192	AY584192 Homo sapi
c 25	161	30.4	22758	8	AP004508	AP004508 Lotus cor
c 26	161	30.4	123943	2	AC134968	AC134968 Mus muscu
c 27	161	30.4	146492	2	AC120206	AC120206 Canis fam
28	160.5	30.3	384	6	AX302503	AX302503 Sequence
29	160.5	30.3	384	9	CR456902	CR456902 Homo sapi
30	160.5	30.3	384	9	HSPCAP15	X79805 H.sapiens m
31	160.5	30.3	384	9	HSU12979	U12979 Human trans
32	160.5	30.3	469	6	CQ694593	CQ694593 Sequence
33	160.5	30.3	500	6	CQ706480	CQ706480 Sequence
34	160.5	30.3	592	6	AX360495	AX360495 Sequence
35	160.5	30.3	601	6	AX360658	AX360658 Sequence
36	160.5	30.3	674	6	AX321946	AX321946 Sequence
37	160.5	30.3	685	6	CQ700744	CQ700744 Sequence
38	160.5	30.3	689	6	AX392710	AX392710 Sequence
39	160.5	30.3	701	6	BD203718	BD203718 Human nuc
40	160.5	30.3	701	6	AX014873	AX014873 Sequence
41	160.5	30.3	793	9	BC018189	BC018189 Homo sapi
42	160.5	30.3	1317	9	BC010537	BC010537 Homo sapi
43	160.5	30.3	1336	9	BC009610	BC009610 Homo sapi
44	160.5	30.3	1347	9	BC022339	BC022339 Homo sapi
45	160.5	30.3	1391	6	CQ493863	CQ493863 Sequence

ALIGNMENTS

RESULT 1
AK072440
LOCUS AK072440 691 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023108D18, full insert sequence.
ACCESSION AK072440
VERSION AK072440.1 GI:32982463
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,

Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 691)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .691
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023108D18"

ORIGIN

Alignment Scores:

Pred. No.:	1.26e-41	Length:	691
Score:	530.00	Matches:	101
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-10-629-953-4 (1-101) x AK072440 (1-691)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 146

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 206

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 266

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATG 326

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 386

Qy        101 Ala 101
              |||
Db        387 GCG 389
```

RESULT 2
AR246100
LOCUS AR246100 289 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1459 from patent US 6476212.
ACCESSION AR246100
VERSION AR246100.1 GI:27293974
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 289)
AUTHORS Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.
TITLE Polynucleotides and polypeptides derived from corn ear
JOURNAL Patent: US 6476212-A 1459 05-NOV-2002;
FEATURES Location/Qualifiers
source 1. .289
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:

Pred. No.:	1.32e-26	Length:	289
Score:	368.00	Matches:	70
Percent Similarity:	86.17%	Conservative:	11
Best Local Similarity:	74.47%	Mismatches:	11
Query Match:	69.43%	Indels:	2
DB:	6	Gaps:	1

US-10-629-953-4 (1-101) x AR246100 (1-289)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||   ||||||   ||||||||||||||||||      ||||||||||||||||||
Db          6 ATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 65

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   ||||||||||||||||||||||:::|||       |||||||||
Db         66 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 125

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
```

```
            :::||||||||||||:::::::|||:::||||||:::|||   |||:::|||   ||||||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGAC 185

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||   ||||||||||||||||||||   ||||||||||   |||
Db        186 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 245

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAla 92
                 :::||||||||||||||||||:::||||||||||||||||
Db        246 CAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCC 287

RESULT 3
AP004876
LOCUS       AP004876               145419 bp    DNA     linear   PLN 28-JUL-2004
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
            PAC clone:P0470G10.
ACCESSION   AP004876
VERSION     AP004876.3  GI:46390400
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
            clone:P0470G10
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 145419)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     On Apr 14, 2004 this sequence version replaced gi:34850224.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'.  A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene
            predicted by a single gene prediction program is also classified as
            a probable 'hypothetical' protein and is included as a
            miscellaneous feature of the sequence.
            The orientation of the sequence is from SP6 to T7 of the PAC clone.
            This sequence of P0470G10 clone has an overlap with OJ1004_H01
            (DDBJ: AP004038) clone at 5' end and with OJ1008_F08 (DDBJ:
            AP004017) clone at 3' end. Detailed information on overlap and
            assembly quality together with annotation of this entry is
            available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1. .145419
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="2"
                     /clone="P0470G10"
     gene            complement(340. .3372)
                     /gene="P0470G10.1"
     misc_feature    complement(340. .3372)
                     /gene="P0470G10.1"
                     /note="probably inactive due to 5' exon missing in CDS
                     pseudogene, receptor protein kinase"
                     /pseudo
     gene            complement(4490. .5523)
                     /gene="P0470G10.2"
     misc_feature    complement(4490. .5523)
                     /gene="P0470G10.2"
                     /note="probably inactive due to 3' exon missing in CDS
                     pseudogene, Epstein-Barr virus EBNA-1-like protein"
                     /pseudo
     gene            complement(7419. .8108)
                     /gene="P0470G10.3"
     misc_feature    complement(7419. .8108)
                     /gene="P0470G10.3"
                     /note="probably inactive due to 3' exon missing in CDS
                     pseudogene, Hcr2-0A"
                     /pseudo
     gene            complement(8546. .10212)
                     /gene="P0470G10.4"
     mRNA            complement(join(<8546. .8551,9081. .9188,10147. .>10212))
                     /gene="P0470G10.4"
                     /note="start and end point are not identified"
     CDS             complement(join(8546. .8551,9081. .9188,10147. .10212))
                     /gene="P0470G10.4"
                     /note="predicted by GENSCAN etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD33295.1"
                     /db_xref="GI:50725764"
                     /translation="MARQVARAVVALVLRIEHKHNLRGGAWRRLILAIQDDVELLLLS
                     LLLVLVGGIPGPHLQ"
     gene            complement(11482. .14658)
                     /gene="P0470G10.5"
     mRNA            complement(join(<11482. .11634,11714. .11941,12441. .12545,
                     12620. .12697,12780. .12854,13509. .13595,13688. .13755,
                     13868. .13955,14055. .14155,14265. .14385,14434. .14479,
                     14585. .>14658))
                     /gene="P0470G10.5"
                     /note="start and end point are not identified"
     CDS             complement(join(11482. .11634,11714. .11941,12441. .12545,
                     12620. .12697,12780. .12854,13509. .13595,13688. .13755,
                     13868. .13955,14055. .14155,14265. .14385,14434. .14479,
                     14585. .14658))
                     /gene="P0470G10.5"
                     /note="contains full-length cDNA(s): AK100016,AK058244"
                     /codon_start=1
                     /product="putative thylakoid-bound ascorbate peroxidase"
                     /protein_id="BAD33296.1"
                     /db_xref="GI:50725765"
                     /translation="MRGTCRRLRRRGAGAPARGRGRWSGYGTPAPSQLSHPYEQVRLG
                     WHDSGTYDKNIEEWPQRGGADGSLRFDAELSHGANAGLINALKLIQPIKDKYPGITYA
                     DLFQLASATAIEEAGGPKIPMKYGRVDVTAAEQCPPEGRLPDAGPRVPADHLREVFYR
                     MGLDDKEIVALSGAHTLGRSRPDRSGWGKPETKYTKDGPGEPGGQSWTVEWLKFDNSY
                     FKDIKEQRDQDLLVLPTDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAHAKLSDLGAK
                     FDPPEGFSLDDEPAVEEKDPEPAPAPAAAPPPPPVEEKKEAEPTPVPVTVGAAVASSP
                     ADDNNGAAPQPEPFVAAKYSYGKKELSDSMKQKIRAEYEGFGGSPDKPLQSNYFLNIM
                     LLIGGLAFLTSLLGS"
     gene            16256. .16777
                     /gene="P0470G10.6"
     mRNA            <16256. .>16777
                     /gene="P0470G10.6"
                     /note="start and end point are not identified"
```

```
    CDS             16256. .16777
                    /gene="P0470G10.6"
                    /note="predicted by GeneMark.hmm etc."
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="BAD33297.1"
                    /db_xref="GI:50725766"
                    /translation="MPKTMSPRREATEAPPPPTSVKVSKRTETRLSPGRAARGVETAQ
                    QRPQEGSDARNANPTKAGLGFHTLSVACISTADTPLLHHHLNLRWRVGTTTPAPPPAS
                    LPMPPPAVPHRVALACTSPASTSASLASCCAGHISMRHAGLRHWHPSLPAPPSVAATS
                    QTRLASLHTGCAS"
    gene            complement(18792. .19088)
                    /gene="P0470G10.7"
    misc_feature    complement(18792. .19088)
                    /gene="P0470G10.7"
                    /note="hypothetical ORF
                    predicted by GlimmerM
                    this category is not included in IRGSP standard"
    gene            complement(20341. .21044)
                    /gene="P0470G10.8"
    mRNA            complement(<20341. .>21044)
                    /gene="P0470G10.8"
                    /note="supported by full-length cDNA(s): AK110958"
    misc_feature    complement(20341. .21044)
                    /gene="P0470G10.8"
                    /note="contains full-length cDNA(s): AK110958
                    non-coding transcript
                    probably inactive due to including stop codon(s) in CDS"
    gene            22104. .25416
                    /gene="P0470G10.9"
    mRNA            join(<22104. .22246,23442. .23556,24206. .24344,24782. .24813,
                    24943. .25029,25192. .>25416)
                    /gene="P0470G10.9"
                    /note="start and end point are not identified"
    CDS             join(22104. .22246,23442. .23556,24206. .24344,24782. .24813,
                    24943. .25029,25192. .25416)
                    /gene="P0470G10.9"
                    /note="contains EST(s): C28326(C60709)"
                    /codon_start=1
                    /product="putative tetratricopeptide repeat
                    (TPR)-containing protein"
                    /protein_id="BAD33298.1"
                    /db_xref="GI:50725767"
                    /translation="MRGFLMHQLLDSSSSDDDDELILAAALIAQHQYDIDNAPRRRGG
                    SVRSDNDETEKEETRVVTTDCPGTKPSDNAESLKHQGNRLAEEGKYHEALGKWEAALT
                    LMPNNAILHEQKAQILLELGDAWRALTAATRATELDPLWPEGWVTLGRAQLNFGEPDS
                    AILSFDKALAIEPDNNEAKSDRETAARLVKKRGQLHSSGLSANKRRFTVGEDSEKEKD
                    DSQMEDVEKCKEDEEKERDDSQLENVEK"
    gene            27139. .29341
                    /gene="P0470G10.10"
    mRNA            join(27139. .27341,27451. .27556,28514. .28536,28983. .29341)
                    /gene="P0470G10.10"
                    /note="supported by full-length cDNA(s): AK072440"
    CDS             join(27225. .27341,27451. .27556,28514. .28536,28983. .29042)
                    /gene="P0470G10.10"
                    /note="contains EST(s): AU082888(C53889),AU068340(C20066)
                    contains full-length cDNA(s): AK072440"
                    /codon_start=1
                    /product="transcriptional coactivator p15 (PC4) family
                    protein-like"
                    /protein_id="BAD33299.1"
                    /db_xref="GI:50725768"
                    /translation="MWRKGNKRFGGGGEPAAKRRAAGDDGPSESADDDIVVAQISKNR
                    RVAVRTWNGKVVVDIREFYEKDGKTLPGRKGIQLPMDQWKILRDNIKAIDEAIKENA"
    gene            complement(30379. .31415)
                    /gene="P0470G10.11"
    mRNA            complement(join(<30379. .30401,31091. .>31415))

Alignment Scores:
Pred. No.:              3.21e-22        Length:         145419
Score:                  352.00          Matches:        75
Percent Similarity:     66.96%          Conservative:   0
Best Local Similarity:  66.96%          Mismatches:     0
Query Match:            66.42%          Indels:         37
DB:                     8               Gaps:           1

US-10-629-953-4 (1-101) x AP004876 (1-145419)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27225 ATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 27284

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGln--- 39
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27285 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGGTG 27344

Qy         39 ------------------------------------------------------------ 39

Db      27345 CTTCCCTAATCCCCGCACTCAAACCTCACTAAAACCCTAGCGCGGCGGGGATCTTCATTT 27404

Qy         40 ----------------------------------------------IleSerLysAsnAr 44
                                                            ||||||||||||||
Db      27405 TCATTTTTTTTTTTGGGGGGGGGGGGGGGGGTTGGTTTGGCCATAGATATCGAAGAACAG 27464

Qy         44 gArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyrGl 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27465 GAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGCGAGTTCTACGA 27524

Qy         64 uLysAspGlyLysThrLeuProGlyArgLysGly 75
              |||||||||||||||||||||||||||||||
Db      27525 GAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGT 27558

RESULT 4
AR246253
LOCUS       AR246253                263 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 1612 from patent US 6476212.
ACCESSION   AR246253
VERSION     AR246253.1  GI:27294127
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 263)
  AUTHORS   Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.
  TITLE     Polynucleotides and polypeptides derived from corn ear
  JOURNAL   Patent: US 6476212-A 1612 05-NOV-2002;
FEATURES             Location/Qualifiers
     source          1. .263
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              6.88e-20        Length:         263
Score:                  297.00          Matches:        60
Percent Similarity:     82.56%          Conservative:   11
Best Local Similarity:  69.77%          Mismatches:     13
Query Match:            56.04%          Indels:         3
DB:                     6               Gaps:           1

US-10-629-953-4 (1-101) x AR246253 (1-263)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||   ||||||     |||||||||||||||     ||||||||||||||||||
Db          6 ATGTGGGGGAAAGGAAAGANGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 65

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   ||||||||||||||||||||||:::|||      |||||||||
Db         66 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 125

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||:::::::|||:::||||||:::|||   |||:::|||      |||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGNCGAC 185

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||||   ||||||||||||||||||||   ||||||||||||   :::
```

```
Db        186 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATC-ATC 244

Qy         79 ProMetAspGlnTrpLys 84
                :::||||||||||||
Db        245 CAATTAGATCAGTGGAAG 262

RESULT 5
AR251557
LOCUS       AR251557                286 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 6916 from patent US 6476212.
ACCESSION   AR251557
VERSION     AR251557.1  GI:27299431
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Lalgudi,R.V., Ito,L.Y. and Sherman,B.K.
  TITLE     Polynucleotides and polypeptides derived from corn ear
  JOURNAL   Patent: US 6476212-A 6916 05-NOV-2002;
FEATURES             Location/Qualifiers
     source          1. .286
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              9.35e-18        Length:         286
Score:                  275.00          Matches:        57
Percent Similarity:     83.54%          Conservative:   9
Best Local Similarity:  72.15%          Mismatches:     10
Query Match:            51.89%          Indels:         3
DB:                     6               Gaps:           1

US-10-629-953-4 (1-101) x AR251557 (1-286)

Qy          1 MetTrpArgLysGlyAsnLysArgPhe-GlyGlyGly------GlyGluProAlaAlaLy 18
              |||      ||||||   |||||||| |||||||||      ||||||||||||||||
Db         50 ATGTNGGGGAAAGGAAAGAAGCGTTTCGGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAA 109

Qy         18 sArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAl 38
              ||||:::||||||   ||||||||||||||||||||||:::|||      ||||||||
Db        110 GCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGC 169

Qy         38 aGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAs 58
              |:::||||||||||||::::::|||:::||||||:::|||   |||:::|||   |||||
Db        170 CGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGA 229

Qy         58 pIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIle 76
              |:::||||||||||||   ||||||||||||||||||||   ||||||||||||
Db        230 CTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA 284

RESULT 6
BT004745
LOCUS       BT004745                324 bp    mRNA    linear   PLN 22-FEB-2003
DEFINITION  Arabidopsis thaliana At5g09250 gene, complete cds.
ACCESSION   BT004745
VERSION     BT004745.1  GI:28466804
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 324)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 324)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-FEB-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J.,
            Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M.,
            Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yamada,K., Yu,G., Yuan,S., Davis,R.W.,  Theologis,A., and
            Ecker,J.R.

            Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .324
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="U23341"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS             1. .324
                     /note="putative transcriptional co-activator (KIWI)"
                     /codon_start=1
                     /product="At5g09250"
                     /protein_id="AAO44011.1"
                     /db_xref="GI:28466805"
                     /translation="MSSRGKRKDEDVRASDDESETHAPAKKVAKPADDSDQSDDIVVC
                     NISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKKGISLSVDQWNTLRNHAEDIEKA
                     LSDLS"
ORIGIN

Alignment Scores:
Pred. No.:              4.43e-15        Length:         324
Score:                  247.50          Matches:        48
Percent Similarity:     70.59%          Conservative:   12
Best Local Similarity:  56.47%          Mismatches:     22
Query Match:            46.70%          Indels:         3
DB:                     8               Gaps:           1

US-10-629-953-4 (1-101) x BT004745 (1-324)

Qy         15 ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAsp 34
              ||||||   |||                ||||||                    ||||||
Db         70 CCGGCTAAGAAAGTCGCAAAGCCAGCAGACGAT---------TCCGACCAGTCTGACGAT 120

Qy         35 IleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLys 54
              |||||||||      ||||||||||||||||||||:::||||||   ||||||||||||
Db        121 ATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAG 180

Qy         55 ValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74
```

```
            :::   :::|||||||||||||||||||   |||||||||||||||||||||||||:::|||
Db        181 ATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAA 240

Qy         75 GlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAsp 94
              ||||||   |||   :::|||||||||      ||||||::::::   :::   |||:::
Db        241 GGTATCTCTCTAAGCGTGGATCAGTGGAACACTCTTCGGAACCACGCAGAGGATATCGAG 300

Qy         95 GluAlaIleLysGlu 99
              :::|||:::   :::
Db        301 AAGGCCCTCTCTGAC 315

RESULT 7
AF053302
LOCUS       AF053302                605 bp    mRNA    linear   PLN 03-JUL-2001
DEFINITION  Arabidopsis thaliana putative transcriptional co-activator (KIWI)
            mRNA, complete cds.
ACCESSION   AF053302
VERSION     AF053302.1  GI:2997683
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Cormack,R.S., Hahlbrock,K. and Somssich,I.E.
  TITLE     Isolation of putative plant transcriptional coactivators using a
            modified two-hybrid system incorporating a GFP reporter gene
  JOURNAL   Plant J. 14 (6), 685-692 (1998)
  MEDLINE   98346011
   PUBMED   9681033
REFERENCE   2  (bases 1 to 605)
  AUTHORS   Cormack,R.S., Hahlbrock,K. and Somssich,I.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-1998) Biochemistry, Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Cologne 50829, Germany
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /map="between KG31 and nga249"
     gene            1. .605
                     /gene="KIWI"
     CDS             97. .420
                     /gene="KIWI"
                     /codon_start=1
                     /product="putative transcriptional co-activator"
                     /protein_id="AAC08574.1"
                     /db_xref="GI:2997684"
                     /translation="MSSRGKRKDEDVRASDDESETHAPAKKVAKPADDSDQSDDIVVC
                     NISKNRRVSVRNWNGKIWIDIREFYVKDGKTLPGKKGISLSVDQWNTLRNHAEDIEKA
                     LSDLS"
ORIGIN

Alignment Scores:
Pred. No.:              8.58e-15        Length:         605
Score:                  247.50          Matches:        48
Percent Similarity:     70.59%          Conservative:   12
Best Local Similarity:  56.47%          Mismatches:     22
Query Match:            46.70%          Indels:         3
DB:                     8               Gaps:           1

US-10-629-953-4 (1-101) x AF053302 (1-605)

Qy         15 ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAsp 34
              ||||||   |||              ||||||                  ||||||
Db        166 CCGGCTAAGAAAGTCGCAAAGCCAGCAGACGAT---------TCCGACCAGTCTGACGAT 216

Qy         35 IleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLys 54
              |||||||||      |||||||||||||||||||:::||||||   ||||||||||||
```

```
Db        217 ATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAG 276

Qy         55 ValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74
              :::   :::|||||||||||||||||   |||||||||||||||||||||:::|||
Db        277 ATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAA 336

Qy         75 GlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAsp 94
              ||||||   |||   :::|||||||||      ||||||::::::   :::   |||:::
Db        337 GGTATCTCTCTAAGCGTGGATCAGTGGAACACTCTTCGGAACCACGCAGAGGATATCGAG 396

Qy         95 GluAlaIleLysGlu 99
              :::|||:::   :::
Db        397 AAGGCCCTCTCTGAC 411

RESULT 8
AB050390
LOCUS       AB050390                913 bp    mRNA    linear   PLN 14-SEP-2001
DEFINITION  Brassica rapa BcKELP mRNA for putative transcriptional coactivator,
            complete cds.
ACCESSION   AB050390
VERSION     AB050390.1  GI:13699909
KEYWORDS    .
SOURCE      Brassica rapa
  ORGANISM  Brassica rapa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Matsushita,Y., Deguchi,M., Youda,M., Nishiguchi,M. and Nyunoya,H.
  TITLE     The tomato mosaic tobamovirus movement protein interacts with a
            putative transcriptional coactivator KELP
  JOURNAL   Mol. Cells 12 (1), 57-66 (2001)
  MEDLINE   21445625
   PUBMED   11561731
REFERENCE   2  (bases 1 to 913)
  AUTHORS   Matsushita,Y. and Nyunoya,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2000) Yasuhiko Matsushita, Tokyo University of
            Agriculture and Technology, Gene Research Center; 3-5-8 Saiwai-cho,
            Fuchu, Tokyo 183-8509, Japan (E-mail:ymatsu@cc.tuat.ac.jp,
            Tel:81-42-367-5855, Fax:81-42-367-5855)
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Brassica rapa"
                     /mol_type="mRNA"
                     /strain="S9 homozygote"
                     /db_xref="taxon:3711"
                     /clone="MIP102 (=BcKELP)"
                     /clone_lib="lambda GEX5 Brassica rapa flower bud cDNA"
                     /note="vector:pGEX-Bc2"
     gene            1. .913
                     /gene="BcKELP"
     CDS             14. .511
                     /gene="BcKELP"
                     /codon_start=1
                     /product="putative transcriptional coactivator"
                     /protein_id="BAB41214.1"
                     /db_xref="GI:13699910"
                     /translation="MEEESKAKIEETVREILKESDMTEMTEFKVRNLASERLGIDLSD
                     KSHKAFVRGIVKSFLEEVESKQQQQQDKEEEEEEEERAKEGNKEFDDDGDLIICRLS
                     DKRRVTIQEFRGKSLVSIREYYKKDGKELPSSKGISLTDEQWSTFKKNIPAIEAAVKK
                     MESRV"
     polyA_signal    869. .874
                     /gene="BcKELP"
                     /evidence=not_experimental
ORIGIN

Alignment Scores:
Pred. No.:              1.18e-08        Length:         913
Score:                  185.00          Matches:        37
Percent Similarity:     56.70%          Conservative:   18
Best Local Similarity:  38.14%          Mismatches:     22
Query Match:            34.91%          Indels:         20
```

```
DB:                    8             Gaps:          2

US-10-629-953-4 (1-101) x AB050390 (1-913)

Qy           3 ArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAla 22
               ::::::|||||||||   |||
Db         263 AAGGAGGGAAACAAAGAGTTT---------------------------------------- 283

Qy          23 GlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLys 42
                 ||||||                 |||   |||:::::::::   ::::::|||
Db         284 ---GACGAT-----------------GACGGCGATCTCATCATTTGCAGGCTGTCGGAT 322

Qy          43 AsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPhe 62
                 |||||||||   ::::::   :::   ||||||   :::|||   ||||||||:::
Db         323 AAGAGGAGAGTGACGATTCAGGAGTTTAGAGGAAAGAGTTTGGTTTCCATCAGAGAGTAT 382

Qy          63 TyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGln 82
               |||:::|||||||||||   ||||||      |||||||||   |||      :::|||
Db         383 TACAAGAAAGACGGCAAAGAGCTTCCTTCTTCTAAAGGAATAAGCTTAACAGACGAACAA 442

Qy          83 TrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
               |||          :::   ||||||   ||||||:::   |||:::|||:::
Db         443 TGGTCAACGTTCAAGAAAAATATTCCAGCTATCGAAGCTGCTGTCAAGAAA 493

RESULT 9
BT005847
LOCUS       BT005847                498 bp    mRNA    linear   PLN 18-MAR-2003
DEFINITION  Arabidopsis thaliana At4g10920 gene, complete cds.
ACCESSION   BT005847
VERSION     BT005847.1  GI:29028805
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 498)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAR-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M.,
            Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and
            Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES              Location/Qualifiers
     source           1. .498
                      /organism="Arabidopsis thaliana"
                      /mol_type="mRNA"
                      /db_xref="taxon:3702"
                      /chromosome="4"
                      /clone="U23574"
                      /ecotype="Columbia"
                      /note="This clone is in pUNI 51"
     CDS              1. .498
                      /note="unknown protein"
                      /codon_start=1
                      /product="At4g10920"
                      /protein_id="AAO64782.1"
                      /db_xref="GI:29028806"
                      /translation="MEKETKEKIEKTVIEILSESDMKEITEFKVRKLASEKLAIDLSE
                      KSHKAFVRSVVEKFLDEERAREYENSQVNKEEEDGDKDCGKGNKEFDDDGDLIICRLS
                      DKRRVTIQEFKGKSLVSIREYYKKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK
                      MESRV"
ORIGIN

Alignment Scores:
Pred. No.:              9.66e-09        Length:         498
Score:                  183.00          Matches:        37
Percent Similarity:     56.25%          Conservative:   17
Best Local Similarity:  38.54%          Mismatches:     22
Query Match:            34.53%          Indels:         20
DB:                     8               Gaps:           2

US-10-629-953-4 (1-101) x BT005847 (1-498)

Qy           4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
               ||||||||||||   |||
Db         253 AAAGGAAACAAAGAGTTT---------------------------------------- 270

Qy          24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
                 |||||                 |||   |||:::::::::   ::::::|||
Db         271 GATGAT-----------------GACGGCGATCTTATCATTTGCAGGTTATCGGATAAG 312

Qy          44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
               |||||||||   ::::::   :::   ||||||   :::|||   ||||||||:::|||
Db         313 AGAAGAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGAGTATTAC 372

Qy          64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83
               :::|||||||||||||   ||||||      |||||||||   |||      :::||||||
Db         373 AAGAAAGATGGCAAAGAACTTCCTACTTCTAAAGGAATAAGCTTAACAGATGAACAATGG 432

Qy          84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                         :::   |||:::   ||||||:::   |||:::|||:::
Db         433 TCAACCTTCAAGAAAAACATGCCAGCCATCGAAAATGCTGTCAAGAAA 480

RESULT 10
AF053303
LOCUS       AF053303                730 bp    mRNA    linear   PLN 03-JUL-2001
DEFINITION  Arabidopsis thaliana putative transcriptional co-activator (KELP)
            mRNA, complete cds.
ACCESSION   AF053303
VERSION     AF053303.1  GI:2997685
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Cormack,R.S., Hahlbrock,K. and Somssich,I.E.
  TITLE     Isolation of putative plant transcriptional coactivators using a
```

```
          modified two-hybrid system incorporating a GFP reporter gene
  JOURNAL   Plant J. 14 (6), 685-692 (1998)
  MEDLINE   98346011
   PUBMED   9681033
REFERENCE   2  (bases 1 to 730)
  AUTHORS   Cormack,R.S., Hahlbrock,K. and Somssich,I.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-1998) Biochemistry, Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Cologne 50829, Germany
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="IV"
                     /map="between m518 and BIO206"
     gene            1. .730
                     /gene="KELP"
     CDS             45. .542
                     /gene="KELP"
                     /codon_start=1
                     /product="putative transcriptional co-activator"
                     /protein_id="AAC08575.1"
                     /db_xref="GI:2997686"
                     /translation="MEKETKEKIEKTVIEILSESDMKEITEFKVRKLASEKLAIDLSE
                     KSHKAFVRSVVEKFLDEERAREYENSQVNKEEEDGDKDCGKGNKEFDDDGDLIICRLS
                     DKRRVTIQEFKGKSLVSIREYYKKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK
                     MESRV"
ORIGIN

Alignment Scores:
Pred. No.:                1.45e-08       Length:          730
Score:                    183.00         Matches:         37
Percent Similarity:       56.25%         Conservative:    17
Best Local Similarity:    38.54%         Mismatches:      22
Query Match:              34.53%         Indels:          20
DB:                       8              Gaps:            2

US-10-629-953-4 (1-101) x AF053303 (1-730)

Qy          4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
              ||||||||||||   |||
Db        297 AAAGGAAACAAAGAGTTT---------------------------------------- 314

Qy         24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
              ||||||                |||   |||:::::::::   :::::::|||
Db        315 GATGAT-----------------GACGGCGATCTTATCATTTGCAGGTTATCGGATAAG 356

Qy         44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||||||||   ::::::   :::   ||||||   :::|||   ||||||||||:::|||
Db        357 AGAAGAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGAGTATTAC 416

Qy         64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83
              :::||||||||||||||   ||||||      |||||||||   |||      :::||||||
Db        417 AAGAAAGATGGCAAAGAACTTCCTACTTCTAAAGGAATAAGCTTAACAGATGAACAATGG 476

Qy         84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                  :::   |||:::   |||||||:::   |||:::|||:::
Db        477 TCAACCTTCAAGAAAAACATGCCAGCCATCGAAAATGCTGTCAAGAAA 524

RESULT 11
AY085896
LOCUS       AY085896                 757 bp    mRNA    linear   PLN 14-APR-2003
DEFINITION  Arabidopsis thaliana clone 19449 mRNA, complete sequence.
ACCESSION   AY085896
VERSION     AY085896.1  GI:21404606
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
  TITLE     Full-length messenger RNA sequences greatly improve genome
            annotation
  JOURNAL   Genome Biol. 3 (6), RESEARCH0029 (2002)
  MEDLINE   22088475
   PUBMED   12093376
REFERENCE   2  (bases 1 to 757)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Full-Length cDNA from Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 757)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="19449"
     CDS             91. .588
                     /codon_start=1
                     /product="putative transcriptional coactivator"
                     /protein_id="AAM63108.1"
                     /db_xref="GI:21554027"
                     /translation="MEKETKEKIEKTVIEILSESDMKEITEFKVRKLASEKLAIDLSE
                     KSHKAFVRSVVEKFLDEERAREYENSQVNKEEEDGDKDCGKGNKEFDDDGDLIICRLS
                     DKRRVTIQEFKGKSLVSIREYYKKDGKELPTSKGISLTDEQWSTFKKNMPAIENAVKK
                     MESRV"
ORIGIN

Alignment Scores:
Pred. No.:                1.51e-08       Length:          757
Score:                    183.00         Matches:         37
Percent Similarity:       56.25%         Conservative:    17
Best Local Similarity:    38.54%         Mismatches:      22
Query Match:              34.53%         Indels:          20
DB:                       8              Gaps:            2

US-10-629-953-4 (1-101) x AY085896 (1-757)

Qy          4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
              ||||||||||||   |||
Db        343 AAAGGAAACAAAGAGTTT---------------------------------------- 360

Qy         24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
              ||||||                |||   |||:::::::::   :::::::|||
Db        361 GATGAT-----------------GACGGCGATCTTATCATTTGCAGGTTATCGGATAAG 402

Qy         44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||||||||   ::::::   :::   ||||||   :::|||   ||||||||||:::|||
Db        403 AGAAGAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGAGTATTAC 462
```

```
Qy         64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83
              :::|||||||||||||   ||||||      |||||||||   |||      :::||||||
Db        463 AAGAAAGATGGCAAAGAACTTCCTACTTCTAAAGGAATAAGCTTAACAGATGAACAATGG 522

Qy         84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                      :::   |||:::   ||||||:::   |||:::|||:::
Db        523 TCAACCTTCAAGAAAAACATGCCAGCCATCGAAAATGCTGTCAAGAAA 570
```

RESULT 12
AK104192

LOCUS	AK104192 923 bp mRNA linear PLN 24-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:006-303-D08, full insert sequence.
ACCESSION	AK104192
VERSION	AK104192.1 GI:32989401
KEYWORDS	FLI_CDNA; oligo capping.
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 923)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE	Direct Submission
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cDNA/ NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M. FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES	Location/Qualifiers
source	1. .923
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="mRNA"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/clone="006-303-D08"

ORIGIN

Alignment Scores:			
Pred. No.:	2.31e-08	Length:	923
Score:	182.00	Matches:	41
Percent Similarity:	62.00%	Conservative:	21
Best Local Similarity:	41.00%	Mismatches:	27
Query Match:	34.34%	Indels:	11
DB:	8	Gaps:	3

US-10-629-953-4 (1-101) x AK104192 (1-923)

```
Qy          3 ArgLysGlyAsnLysArgPheGlyGlyGlyGly-GluPro------AlaAlaLysArgAr 20
              |||   |||       |||   ||||||||||| |||                  :::||
Db        333 CGAGGAGGAAGAGGAAGAGGAGGAGGAGGAGGAGGAGCGGGGAGGTGGAACGAAGAG 392

Qy         20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIl 40
              |     |||:::|||                        |||:::::::::   :::::
Db        393 GGAGTACGATGACGAGGGA--------------------GACCTCATCCTCTGCCGCCT 431

Qy         40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
              :|||      |||||||||   ::::::   :::   ||||||   ::::::   |||||
Db        432 TTCGGCCAGGAGGAGGGTGACTCTTCAGGAGTTCAAGGGCAAGACGCTTCTGTCCATCCG 491

Qy         60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMe 80
              ||||:::|||   ||||||||||||   ||||||      |||||||||   |||   ::
Db        492 AGAGTACTACTTCAAGGACGGCAAGGAGCTGCCCGCC---AAAGGGATAAGTTTGACAGT 548

Qy         80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              ::::||||||:::      ||||||::::::   ||||||::::::|||||||||:::
Db        549 TGAGCAGTGGGAAGCATTTCGTGATTCTGTGCCAGCAATAGAGGATGCCATTAAAAAG 606
```

RESULT 13
AK070545

LOCUS	AK070545 945 bp mRNA linear PLN 24-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J023055M24, full insert sequence.
ACCESSION	AK070545

VERSION AK070545.1 GI:32980569
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 945)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
FEATURES Location/Qualifiers
source 1. .945
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023055M24"
ORIGIN

Alignment Scores:

Pred. No.:	2.37e-08	Length:	945
Score:	182.00	Matches:	41
Percent Similarity:	62.00%	Conservative:	21
Best Local Similarity:	41.00%	Mismatches:	27
Query Match:	34.34%	Indels:	11
DB:	8	Gaps:	3

US-10-629-953-4 (1-101) x AK070545 (1-945)

```
Qy          3 ArgLysGlyAsnLysArgPheGlyGlyGlyGly-GluPro------AlaAlaLysArgAr 20
              |||   |||      |||   |||||||||||| |||                ::::||
Db        280 CGAGGAGGAAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGCGGGGAGGTGGAACGAAGAG 339

Qy         20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIl 40
              |   |||:::|||                        ||||:::::::::   :::::
Db        340 GGAGTACGATGACGAGGGA--------------------GACCTCATCCTCTGCCGCCT 378

Qy         40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
              :|||      |||||||||   ::::::   :::   ||||||   ::::::   |||||
Db        379 TTCGGCCAGGAGGAGGGTGACTCTTCAGGAGTTCAAGGGCAAGACGCTTCTGTCCATCCG 438

Qy         60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMe 80
              ||||:::|||   ||||||||||||   ||||||      |||||||||   |||   ::
Db        439 AGAGTACTACTTCAAGGACGGCAAGGAGCTGCCCGCC---AAAGGGATAAGTTTGACAGT 495

Qy         80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              ::::||||||:::      ||||||::::::   ||||||::::::|||||||||:::
Db        496 TGAGCAGTGGGAAGCATTTCGTGATTCTGTGCCAGCAATAGAGGATGCCATTAAAAAG 553
```

RESULT 14
AK059936
LOCUS AK059936 1141 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:006-210-D10, full insert sequence.
ACCESSION AK059936
VERSION AK059936.1 GI:32969954
KEYWORDS FLI_CDNA; oligo-capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE 2 (bases 1 to 1141)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Direct Submission

JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers

source 1. .1141
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="006-210-D10"

ORIGIN

Alignment Scores:

Pred. No.:	2.89e-08	Length:	1141
Score:	182.00	Matches:	41
Percent Similarity:	62.00%	Conservative:	21
Best Local Similarity:	41.00%	Mismatches:	27
Query Match:	34.34%	Indels:	11
DB:	8	Gaps:	3

US-10-629-953-4 (1-101) x AK059936 (1-1141)

```
Qy          3 ArgLysGlyAsnLysArgPheGlyGlyGlyGly-GluPro------AlaAlaLysArgAr 20
              |||  |||       |||  ||||||||||| |||                   :::||
Db        358 CGAGGAGGAAGAGGAAGAGGAGGAGGAGGAGGAGGAGCGGGGAGGTGGAACGAAGAG 417

Qy         20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIl 40
              |    |||:::|||                  ||:::::::::   :::::
Db        418 GGAGTACGATGACGAGGGA--------------------GACCTCATCCTCTGCCGCCT 456

Qy         40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
              :|||      |||||||||   ::::::   :::   ||||||   ::::::   |||||
Db        457 TTCGGCCAGGAGGAGGGTGACTCTTCAGGAGTTCAAGGGCAAGACGCTTCTGTCCATCCG 516

Qy         60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMe 80
              ||||:::|||   ||||||||||   ||||||      |||||||||   |||   ::
Db        517 AGAGTACTACTTCAAGGACGGCAAGGAGCTGCCCGCC---AAAGGGATAAGTTTGACAGT 573

Qy         80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              ::::||||||:::      |||||::::::   ||||||:::::|||||||||:::
Db        574 TGAGCAGTGGGAAGCATTTCGTGATTCTGTGCCAGCAATAGAGGATGCCATTAAAAAG 631
```

RESULT 15
ATT2K12/c
LOCUS ATT2K12 82896 bp DNA linear PLN 03-APR-2001
DEFINITION Arabidopsis thaliana DNA chromosome 5, BAC clone T2K12 (ESSA project).
ACCESSION AL590346
VERSION AL590346.1 GI:13548323
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 82896)
AUTHORS Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W., Rudd,S., Lemcke,K. and Mayer,K.F.X.
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 82896)
AUTHORS EU Arabidopsis sequencing,project.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) MIPS, at GSF/IBI, Ingolstaedter Landstr. 1, 85764 Neuherberg, FRG, E-mail: schoof@gsf.de,kmayer@gsf.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk

COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://mips.gsf.de/proj/thal/.

FEATURES Location/Qualifiers

source 1. .82896
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="5"

misc_feature 1. .14902

```
                /note="overlap to PAC MAH20, EMBL:AB006697; please refer
                to this clone for annotation"
gene            15981. .19300
                /gene="At5g08660"
CDS             join(15981. .16289,16433. .16746,16833. .16925,17020. .17110,
                17305. .17398,17478. .17575,17767. .17836,18028. .18131,
                18231. .18347,18458. .18534,18636. .18680,18763. .19300)
                /gene="At5g08660"
                /note="similarity to predicted proteins, Arabidopsis
                thaliana"
                /codon_start=1
                /product="putative protein"
                /protein_id="CAC35871.1"
                /db_xref="GI:13548324"
                /db_xref="TrEMBL:Q9C5B1"
                /translation="MGSFCSKSLGINFGSEYSGSSVADDGREPDFGHSQPNGQTSLIV
                PGMRQLMVKDVKEQNQLKDVFSFRREAEDNFYDGIPTYTMAPSQKIRSAKSTQTAVS
                KVTEASKLLGKAGLGRAKDVLDTLGSSMTDLSSGGFTSGVATKGNELGILAFEVANTI
                VKSSNLIESLSKRNIEHLKGTILYSEGVQNLVSNDFDELLRLVAADKRQELQVFSGEV
                VRFGNRSKDFQWHNLQRYFDRISKELTPQRQLKEDAVLVVDQLMVLVQYTAELYQELQ
                VLYRLEKDYEQKRREEENSANSSKGDGLAILKTELKAQRKVVKSLKKKSLWSRGFEEV
                MEKLVDIVHFLLLEIHNIFGGADDQPSKKGAAEYDKRLGPAGLALHYANIIVQIDTLV
                ARASSITSNARDSLYQSLPPGIKLALRSKIKSFNVDKELSVTQIKDEMERTLHWLVPV
                AGNTTKAHHGFGWVGEWANTGTDFTSKPSGGDILRIETLYHASKEKTEIYILGQIIWL
                QHLVTKAKSDARGGPRLSSIKSPLDTTNQQLISEPLSVPIVTDEEQKMLQEASKRKRT
                PCVSKSQDFDSEYSRARKCDPLSKSSEYFRGVRRSKSAAVKRYSSGFPLLDFAIDKEK
                VLDVIDRVDVPRDYKALLKEGSLSF"
exon            15981. .16289
                /gene="At5g08660"
                /number=1
intron          16290. .16432
                /gene="At5g08660"
                /number=1
exon            16433. .16746
                /gene="At5g08660"
                /number=2
intron          16747. .16832
                /gene="At5g08660"
                /number=2
exon            16833. .16925
                /gene="At5g08660"
                /number=3
intron          16926. .17019
                /gene="At5g08660"
                /number=3
exon            17020. .17110
                /gene="At5g08660"
                /number=4
intron          17111. .17304
                /gene="At5g08660"
                /number=4
exon            17305. .17398
                /gene="At5g08660"
                /number=5
intron          17399. .17477
                /gene="At5g08660"
                /number=5
exon            17478. .17575
                /gene="At5g08660"
                /number=6
intron          17576. .17766
                /gene="At5g08660"
                /number=6
exon            17767. .17836
                /gene="At5g08660"
                /number=7
intron          17837. .18027
                /gene="At5g08660"
                /number=7
exon            18028. .18131
                /gene="At5g08660"
                /number=8
intron          18132. .18230
                /gene="At5g08660"
                /number=8
exon            18231. .18347
                /gene="At5g08660"
                /number=9
intron          18348. .18457
                /gene="At5g08660"
                /number=9
exon            18458. .18534
                /gene="At5g08660"
                /number=10
intron          18535. .18635
                /gene="At5g08660"
                /number=10
exon            18636. .18680
                /gene="At5g08660"
                /number=11
intron          18681. .18762
                /gene="At5g08660"
                /number=11
exon            18763. .19300
                /gene="At5g08660"
                /number=12
gene            19850. .22604
                /gene="At5g08670"
gene            complement(join(19850. .19963,20204. .20515,20610. .20834,
                20918. .21076,21177. .21280,21360. .21477,21569. .21657,
                21738. .21813,22131. .22604))
                /gene="At5g08670"
CDS             complement(join(19850. .19963,20204. .20515,20610. .20834,
                20918. .21076,21177. .21280,21360. .21477,21569. .21657,
                21738. .21813,22131. .22604))
                /gene="At5g08670"
                /note="strong similarity to H+-transporting ATP synthase,
                mitochondrial, rubber tree, PIR:S20504
                Contains ATP/GTP-binding site motif A (P-loop)
                AA231-238;ATP synthase alpha and beta subunits signature
                AA423-432"
                /codon_start=1
                /product="H+-transporting ATP synthase beta chain
                (mitochondrial)-like protein"
                /protein_id="CAC35872.1"
                /db_xref="GI:13548325"
                /db_xref="GOA:P83483"
                /db_xref="Swiss-Prot:P83483"
                /translation="MASRRVLSSLLRSSSGRSAAKLGNRNPRLPSPSPARHAAPCSYL
                LGRVAEYATSSPASSAAPSSAPAKDEGKKTYDYGGKGAIGRVCQVIGAIVDVRFEDQE
                GLPPIMTSLEVQDHPTRLVLEVSHHLGQNVVRTIAMDGTEGLVRGRKVLNTGAPITVP
                VGRATLGRIMNVLGEPIDERGEIKTEHYLPIHRDAPALVDLATGQEILATGIKVVDLL
                APYQRGGKIGLFGGAGVGKTVLIMELINNVAKAHGGFSVFAGVGERTREGNDLYREMI
                ESGVIKLGEKQSESKCALVYGQMNEPPGARARVGLTGLTVAEYFRDAEGQDVLLFIDN
                IFRFTQANSEVSALLGRIPSAVGYQPTLASDLGALQERITTTKKGSITSVQAIYVPAD
                DLTDPAPATTFAHLDATTVLSRQISELGIYPAVDPLDSTSRMLSPHILGEEHYNTARG
                VQKVLQNYKNLQDIIAILGMDELSEDDKLTVARARKIQRFLSQPFHVAEIFTGAPGKY
                VDLKENINSFQGLLDGKYDDLSEQSFYMVGGIDEVVAKAEKIAKESAA"
exon            complement(19850. .19963)
                /gene="At5g08670"
                /number=1
intron          complement(19964. .20203)
                /gene="At5g08670"
                /number=1
exon            complement(20204. .20515)
                /gene="At5g08670"
                /number=2
intron          complement(20516. .20609)
                /gene="At5g08670"
                /number=2
exon            complement(20610. .20834)
                /gene="At5g08670"
                /number=3
intron          complement(20835. .20917)
                /gene="At5g08670"
                /number=3
```

```
     exon            complement(20918. .21076)
                     /gene="At5g08670"
                     /number=4
     intron          complement(21077. .21176)
                     /gene="At5g08670"
                     /number=4
     exon            complement(21177. .21280)
                     /gene="At5g08670"
                     /number=5
     intron          complement(21281. .21359)
                     /gene="At5g08670"
                     /number=5
     exon            complement(21360. .21477)
                     /gene="At5g08670"
                     /number=6
     intron          complement(21478. .21568)
                     /gene="At5g08670"
                     /number=6
     exon            complement(21569. .21657)
                     /gene="At5g08670"

Alignment Scores:
Pred. No.:              1.57e-05        Length:         82896
Score:                  174.00          Matches:        31
Percent Similarity:     90.24%          Conservative:   6
Best Local Similarity:  75.61%          Mismatches:     4
Query Match:            32.83%          Indels:         0
DB:                     8               Gaps:           0

US-10-629-953-4 (1-101) x ATT2K12 (1-82896)

Qy         35 IleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLys 54
              :::|||:::   |||||||||||||||||||||:::||||||   ||||||||||||
Db      77551 TTGGTTATACATCAGATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAG 77492

Qy         55 ValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74
              :::   :::|||||||||||||||||   |||||||||||||||||||||:::|||
Db      77491 ATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAA 77432

Qy         75 Gly 75
              |||
Db      77431 GGT 77429


Search completed: January 12, 2005, 02:38:40
Job time : 3329 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 11, 2005, 23:18:54 ; Search time 412 Seconds
(without alignments)
1286.873 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10629953/runat_10012005_172534_9407/app_query.fasta_1.263
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10629953_@CGN_1_1_470_@runat_10012005_172534_9407 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	530	100.0	632	3	AAZ50412	Aaz50412 Rice tran
2	398	75.1	649	3	AAZ50411	Aaz50411 Corn tran
3	389.5	73.5	460	3	AAZ50414	Aaz50414 Wheat tra
4	368	69.4	289	10	ABX82999	Abx82999 Corn ear-
5	304.5	57.5	310	6	ABL70693	Abl70693 Corn tass
6	297	56.0	263	10	ABX83152	Abx83152 Corn ear-
7	275	51.9	286	10	ABX88456	Abx88456 Corn ear-
8	269	50.8	672	3	AAZ50413	Aaz50413 Soybean t
9	247.5	46.7	471	3	AAC38126	Aac38126 Arabidops
10	206.5	39.0	516	3	AAZ50416	Aaz50416 Vernonia
11	194.5	36.7	1063	3	AAC32760	Aac32760 Arabidops
12	185	34.9	913	8	ACF03526	Acf03526 Brassica
13	184.5	34.8	740	3	AAZ50419	Aaz50419 Soybean t
c 14	183	34.5	496	10	ADE81749	Ade81749 Arabidops
15	183	34.5	757	3	AAC50198	Aac50198 Arabidops
16	183	34.5	757	3	AAC34846	Aac34846 Arabidops
17	168.5	31.8	1089	3	AAZ50417	Aaz50417 Corn tran
18	167.5	31.6	498	3	AAZ50415	Aaz50415 Marigold
19	166	31.3	615	10	ADB57912	Adb57912 Toxicity-
20	166	31.3	615	10	ADB52411	Adb52411 Primary r
21	166	31.3	939	3	AAZ50418	Aaz50418 Corn tran
22	163	30.8	483	9	ACH38105	Ach38105 Human end
23	160.5	30.3	384	5	ABA83091	Aba83091 Polymeras
24	160.5	30.3	384	8	ACC50254	Acc50254 Breast ca
25	160.5	30.3	384	10	ADF76607	Adf76607 Novel hum
26	160.5	30.3	443	8	ABZ56521	Abz56521 Aspergill
27	160.5	30.3	460	9	ACH47965	Ach47965 Human lun
28	160.5	30.3	592	6	ABK16067	Abk16067 Human lun
29	160.5	30.3	592	10	ADB95330	Adb95330 Human lun
30	160.5	30.3	601	6	ABK16230	Abk16230 Human lun
31	160.5	30.3	601	10	ADB95493	Adb95493 Human lun
32	160.5	30.3	674	4	AAD23497	Aad23497 Human lun
33	160.5	30.3	674	10	ADD66785	Add66785 Human lun
34	160.5	30.3	674	10	ADE88039	Ade88039 Human lun
35	160.5	30.3	689	6	ABL42297	Abl42297 Pancreati
36	160.5	30.3	701	2	AAZ77517	Aaz77517 Human ova
37	160.5	30.3	1336	12	ADL91493	Adl91493 Human imm
38	160.5	30.3	1336	12	ADN04675	Adn04675 Antipsori
39	160.5	30.3	1391	5	ABV25741	Abv25741 Human pro
40	160.5	30.3	3527	3	AAC99092	Aac99092 Human pan
c 41	160	30.2	622	6	ABK83518	Abk83518 Human cDN
42	157.5	29.7	432	9	ACH19785	Ach19785 Human adu
43	154.5	29.2	493	5	AAS91456	Aas91456 DNA encod
44	153.5	29.0	662	3	AAC10376	Aac10376 Human sec
c 45	147.5	27.8	547	12	ADP03062	Adp03062 Human hou

ALIGNMENTS

RESULT 1
AAZ50412
ID AAZ50412 standard; cDNA; 632 BP.
XX
AC AAZ50412;
XX
DT 18-MAY-2000 (first entry)
XX
DE Rice transcription coactivator PC4(P15) type 1 cDNA.
XX
KW Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW rice; signal mediator; activator; Gal4/VP16; transgenic plant;
KW general transcription factor; GTF; transcription initiation complex;
KW immunological screening; detection; marker; clone rr1.pk0003.a12; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 46. .351
FT /*tag= a
FT /product= "Rice PC4(P15) type 1 transcription
FT coactivator"
FT /note= "Derived from clone rr1.pk0003.a12"
XX
PN WO200005377-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016479.
XX

```
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44879.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 9; Page 33; 43pp; English.
XX
CC   The present sequence is the cDNA encoding rice Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone rr1.pk0003.a12, obtained from rr1 cDNA library prepared using
CC   rice root two week old developing seedlings. PC4 molecules functions as a
CC   signal mediator between activators like Gal4/VP16 and general
CC   transcription factors (GTFs) in a transcription initiation complex. The
CC   expression levels of PC4 can be manipulated and the functional properties
CC   of specific transcriptional activators can be modulated. The PC4 DNA can
CC   be used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 632 BP; 165 A; 139 C; 187 G; 141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.69e-54      Length:         632
Score:                  530.00        Matches:        101
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     3             Gaps:           0

US-10-629-953-4 (1-101) x AAZ50412 (1-632)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 105

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 165

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 225

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCAATG 285

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 345

Qy        101 Ala 101
              |||
Db        346 GCG 348

RESULT 2
AAZ50411
ID   AAZ50411 standard; cDNA; 649 BP.
XX
AC   AAZ50411;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Corn transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   corn; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone cca.pk0020.d2; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             33. .344
FT                   /*tag=  a
FT                   /product= "Corn PC4(P15) type 1 transcription
FT                   coactivator"
FT                   /note= "Derived from clone cca.pk0020.d2"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44878.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 32; 43pp; English.
XX
CC   The present sequence is the cDNA encoding corn Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from the corn callus type II tissue (undifferentiated) from the cDNA
CC   clone cca.pk0020.d2. PC4 molecules functions as a signal mediator between
CC   activators like Gal4/VP16 and general transcription factors (GTFs) in a
CC   transcription initiation complex. The expression levels of PC4 can be
CC   manipulated and the functional properties of specific transcriptional
CC   activators can be modulated. The PC4 DNA can be used to create transgenic
CC   plants with altered PC4 levels, that would affect the level of
CC   transcription of specific genes in the plant. It is also used for
CC   immunological screening of cDNA libraries and to raise specific
CC   antibodies for detection. The DNA is used as probes and primers, for
CC   genetic and physical mapping of genes and as markers for traits linked to
CC   those genes
XX
SQ   Sequence 649 BP; 167 A; 147 C; 182 G; 153 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.04e-38      Length:         649
Score:                  398.00        Matches:        76
Percent Similarity:     87.25%        Conservative:   13
Best Local Similarity:  74.51%        Mismatches:     11
Query Match:            75.09%        Indels:         2
DB:                     3             Gaps:           1

US-10-629-953-4 (1-101) x AAZ50411 (1-649)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||   ||||||   ||||||||||||||||||      ||||||||||||||||||
Db         33 ATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 92

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::|||||||   ||||||||||||||||||||||:::|||      |||||||||
Db         93 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 152
```

```
Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::||||||||||||:::::|||:::|||||:::|||   |||:::|||   ||||||
Db        153 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGAC 212

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::||||||||||||   ||||||||||||||||||||   ||||||||||   |||
Db        213 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 272

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98
                 :::||||||||||||||||:::||||||||||||||||:::||||||||:::
Db        273 CAATTAGATCAGTGGAAGATACTGAAGGACAACATCAAAGCCATAAATGAGGCAATCGAG 332

Qy         99 GluAsn 100
              ||||||
Db        333 GAAAAT 338
```

```
RESULT 3
AAZ50414
ID   AAZ50414 standard; cDNA; 460 BP.
XX
AC   AAZ50414;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Wheat transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   wheat; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone wdk2c.pk015.g20; ss.
XX
OS   Triticum aestivum.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .293
FT                   /*tag=  a
FT                   /product= "Wheat PC4(P15) type 1 transcription
FT                   coactivator"
FT                   /note= "Derived from clone wdk2c.pk015.g20"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44881.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 35; 43pp; English.
XX
CC   The present sequence is the cDNA encoding wheat Positive Cofactor 4 (PC4)
CC   transcription coactivator, designated as PC4(P15) type 1. It is isolated
CC   from clone wdk2c.pk015.g20, obtained from wdk2c cDNA library, prepared
CC   from wheat developing kernels, 7 days after anthesis. PC4 molecules
CC   functions as a signal mediator between activators like Gal4/VP16 and
CC   general transcription factors (GTFs) in a transcription initiation
CC   complex. The expression levels of PC4 can be manipulated and the
CC   functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 460 BP; 117 A; 110 C; 139 G; 87 T; 0 U; 7 Other;
```

```
Alignment Scores:
Pred. No.:              2.72e-37      Length:          460
Score:                  389.50        Matches:         76
Percent Similarity:     87.37%        Conservative:    7
Best Local Similarity:  80.00%        Mismatches:      9
Query Match:            73.49%        Indels:          3
DB:                     3             Gaps:            1

US-10-629-953-4 (1-101) x AAZ50414 (1-460)

Qy          9 PheGlyGlyGly---------GlyGluProAlaAlaLysArgArgAlaAlaGlyAspAsp 25
              ||||||||||||         |||:::|||   |||||||||:::|||||||||   |||
Db          3 TTCGGCGGCGGCGGCGGCCGCGGCCAGCCGCCGGCCAAGCGCCAGGCGGCCGGGAAGGAC 62

Qy         26 GlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
              ||||||||||||       ||||||   ||||||||||||||||||||||||||:::|||
Db         63 GGTCCCTCCGAGGAAACCGACGACGGCATCGTCGTTGCGCAGATATCGAAGAACAAGAGG 122

Qy         46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyrGluLys 65
              ||||||||||||   |||||||||||:::||||||:::|||||||||||||||||||||
Db        123 GTGGCCGTGAGGAACTGGAACGGGAAGGTCATGGTCGACATGCGCGAGTTCTACGAAAAG 182

Qy         66 AspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIle 85
              ||||||||:::||||||  ||||||||||||   |||   ||||||||||||||||||
Db        183 GACGGCAAGAGCCTCCCGACCCGCAAAGGTATATCGCTCTCAATGGATCAGTGGAAAATA 242

Qy         86 LeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||:::||||||||||||||||||||||||||||
Db        243 CTGAGGGACAACATCGAAGCTATAGACGAGGCCATCAAGGAGAAC 287
```

```
RESULT 4
ABX82999
ID   ABX82999 standard; cDNA; 289 BP.
XX
AC   ABX82999;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #1459.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
```

```
XX
PS   Example; SEQ ID NO 1459; 390pp; English.
XX
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 289 BP; 79 A; 71 C; 92 G; 47 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.63e-35        Length:         289
Score:                  368.00          Matches:        70
Percent Similarity:     86.17%          Conservative:   11
Best Local Similarity:  74.47%          Mismatches:     11
Query Match:            69.43%          Indels:         2
DB:                     10              Gaps:           1

US-10-629-953-4 (1-101) x ABX82999 (1-289)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||  ||||||  |||||||||||||||||      ||||||||||||||||||
Db          6 ATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 65

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||  ||||||||||||||||||||||:::|||      |||||||||
Db         66 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 125

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||:::::::|||:::||||||:::|||   |||:::|||   ||||||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGAC 185

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::||||||||||||   ||||||||||||||||||   ||||||||||   |||
Db        186 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 245

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAla 92
                 :::||||||||||||||||||:::|||||||||||||||
Db        246 CAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCC 287

RESULT 5
ABL70693
ID   ABL70693 standard; cDNA; 310 BP.
XX
AC   ABL70693;
XX
DT   14-MAY-2002  (first entry)
XX
DE   Corn tassel-derived polynucleotide (cdps) SEQ ID NO:67.
XX
KW   Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW   inheritance; characteristic; growth; development; disease resistance;
KW   environmental adaptability; quality; yield; molecular marker;
KW   multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS   Zea mays.
XX
PN   US2001051335-A1.
XX
PD   13-DEC-2001.
XX
PF   16-APR-1999;   99US-00294093.
XX
PR   21-APR-1998;   98US-0082567P.
XX
PA   (LALG/) LALGUDI R V.
PA   (ITOL/) ITO L Y.
PA   (SHER/) SHERMAN B K.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2002-163647/21.
XX
PT   Novel purified corn tassel-derived polynucleotide useful for determining
PT   altered gene expression, to recover regulatory elements and to follow
PT   inheritance of desirable characteristics through hybrid breeding
PT   programs.
XX
PS   Claim 1; SEQ ID NO 67; 201pp; English.
XX
CC   The present sequence describes a purified corn tassel-derived
CC   polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC   selected from those given in ABL70627 to ABL76833. The cdps sequences
CC   encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC   can be used for determining altered gene expression, to recover
CC   regulatory elements and to follow inheritance of desirable
CC   characteristics through hybrid breeding programs. (I) are also useful in
CC   the evaluation, and alteration of desired characteristics associated with
CC   growth and development, disease resistance, environmental adaptability,
CC   quality and yield, and as molecular markers for studying inheritance of
CC   multigene traits in a plant breeding program. (I) can be used to produce
CC   a tassel-specific profile of gene transcription, a transcript image, to
CC   clone regulatory elements for use in transformation vectors, to express a
CC   polypeptide, to identify, isolate or extend identical or related corn
CC   tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC   hybridisation or amplification technologies, as query sequences to
CC   determine homology of known sequences, as probe for use in Southern or
CC   Northern hybridisation, and to identify the presence of and/or to
CC   determine the degree of similarity between two (or more) nucleic acid
CC   sequences
XX
SQ   Sequence 310 BP; 84 A; 77 C; 96 G; 50 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              2.36e-27        Length:         310
Score:                  304.50          Matches:        65
Percent Similarity:     81.44%          Conservative:   14
Best Local Similarity:  67.01%          Mismatches:     13
Query Match:            57.45%          Indels:         5
DB:                     6               Gaps:           2

US-10-629-953-4 (1-101) x ABL70693 (1-310)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              |||      :::|||   ||||||||||||||||||      ||||||||||||||||||
Db         26 ATGTNGGGGAGGGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGCGAGCCCGCGGCCAAG 85

Qy         19 ArgArgAlaAla-GlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAl 38
              |||:::|||||| ||||||||||||||||||||||||||:::|||      ||||||||
Db         86 CGCCAGGCCGCGAGGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGC 145

Qy         38 aGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAs 58
              |:::||||||||||||:::::|||:::||||||:::|||   |||:::|||   |||||
Db        146 CGAGATATCGAAGAACAAGAAGGTATCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGA 205

Qy         58 pIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGln-L 78
```

```
            |:::|||||||||||||   ||||||||||||||||||||   ||||||   :::    :
Db        206 CTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAA---GTATATCA 262

Qy         78 euProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIle 93
              ::|||:::     ||||||||||||:::|||   ||||||||||||
Db        263 TTCCAATTAGATCATGGAAGATATTGAAGGACNACATCAAAGCCATA 309

RESULT 6
ABX83152
ID   ABX83152 standard; cDNA; 263 BP.
XX
AC   ABX83152;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #1612.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
XX
PS   Example; SEQ ID NO 1612; 390pp; English.
XX
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 263 BP; 67 A; 65 C; 88 G; 41 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              1.51e-26        Length:         263
Score:                  297.00          Matches:        60
Percent Similarity:     82.56%          Conservative:   11
Best Local Similarity:  69.77%          Mismatches:     13
Query Match:            56.04%          Indels:         3
DB:                     10              Gaps:           1

US-10-629-953-4 (1-101) x ABX83152 (1-263)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||   ||||||      ||||||||||||||      ||||||||||||||||||
Db          6 ATGTGGGGGAAAGGAAAGANGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 65

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   ||||||||||||||||||||||||:::|||      |||||||||
Db         66 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 125

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||:::::|||:::||||||:::|||   |||:::|||      |||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGNCGAC 185

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||   ||||||||||||||||||||   ||||||||||||   :::
Db        186 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATC-ATC 244

Qy         79 ProMetAspGlnTrpLys 84
                 :::|||||||||||||
Db        245 CAATTAGATCAGTGGAAG 262

RESULT 7
ABX88456
ID   ABX88456 standard; cDNA; 286 BP.
XX
AC   ABX88456;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #6916.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
XX
PS   Example; SEQ ID NO 6916; 390pp; English.
XX
```

```
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 286 BP; 69 A; 77 C; 92 G; 47 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              7.12e-24        Length:         286
Score:                  275.00          Matches:        57
Percent Similarity:     83.54%          Conservative:   9
Best Local Similarity:  72.15%          Mismatches:     10
Query Match:            51.89%          Indels:         3
DB:                     10              Gaps:           1

US-10-629-953-4 (1-101) x ABX88456 (1-286)

Qy          1 MetTrpArgLysGlyAsnLysArgPhe-GlyGlyGly------GlyGluProAlaAlaLy 18
              |||      ||||||   |||||||| |||||||||      ||||||||||||||||
Db         50 ATGTNGGGGAAAGGAAAGAAGCGTTTCGGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAA 109

Qy         18 sArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAl 38
              ||||:::||||||   |||||||||||||||||||||||:::|||      ||||||||
Db        110 GCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGC 169

Qy         38 aGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAs 58
              |:::|||||||||||:::::||:::|||:::|||||   |||:::|||   |||||
Db        170 CGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGA 229

Qy         58 pIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIle 76
              |:::||||||||||||   ||||||||||||||||||||   ||||||||||||
Db        230 CTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA 284

RESULT 8
AAZ50413
ID   AAZ50413 standard; cDNA; 672 BP.
XX
AC   AAZ50413;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Soybean transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   soybean; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone sfl1.pk0008.a4; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             37. .354
FT                   /*tag= a
FT                   /product= "Soybean PC4(P15) type I transcription
FT                   coactivator"
FT                   /note= "Derived from clone sfl1.pk0008.a4"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44880.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 33-34; 43pp; English.
XX
CC   The present sequence is the cDNA encoding soybean Positive Cofactor 4
CC   (PC4) transcription coactivator, designated as PC4(P15) type 1. It is
CC   isolated from clone sfl1.pk0008.a4, obtained from sfl1 cDNA library,
CC   prepared from soybean immature flowers. PC4 molecules functions as a
CC   signal mediator between activators like Gal4/VP16 and general
CC   transcription factors (GTFs) in a transcription initiation complex. The
CC   expression levels of PC4 can be manipulated and the functional properties
CC   of specific transcriptional activators can be modulated. The PC4 DNA can
CC   be used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 672 BP; 218 A; 113 C; 154 G; 187 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.08e-22        Length:         672
Score:                  269.00          Matches:        56
Percent Similarity:     66.99%          Conservative:   13
Best Local Similarity:  54.37%          Mismatches:     14
Query Match:            50.75%          Indels:         20
DB:                     3               Gaps:           3

US-10-629-953-4 (1-101) x AAZ50413 (1-672)

Qy         17 AlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSer----------------- 30
              ||||||||||||          |||||||||   |||::::::
Db         49 GCGAAGAGGAGA------GATGACGACGGTGCTTCCGACGCCGACTCTGAAGGCCACGCG 102

Qy         31 --------------------------AlaAspAspAsp---------IleValValAla 38
                                        :::|||||||||         |||   |||
Db        103 CCGCCCAAGAAGTCCCTCAAGAAGGATTCCGATGACGATCCCGACTCTATTACCGTTTGC 162

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||||||||||||||||||||||||   |||   |||   ::::::||||||
Db        163 GAGATTTCGAAGAACAGGAGGGTTGCCGTGAGGAACTGGAAAGGCAGCATTATGGTTGAC 222

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              ||||||||||||||   ||||||||||||   ||||||||||||||||||||   |||
Db        223 ATTCGCGAGTTTTACGTCAAAGATGGCAAGCAATTGCCTGGCAGAAAAGGTATCTCTTTG 282

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98
              ||||||||||||   :::||||||:::::::::::::   ||||||:::||||||
Db        283 ACCATGGATCAGTGGAATGTGCTTCGTAATCATGTTGAAGAAATTGACAAGGCAATTAAT 342

Qy         99 GluAsnAla 101
```

```
                    ||||||:::
Db          343 GAGAATTCT 351

RESULT 9
AAC38126
ID   AAC38126 standard; DNA; 471 BP.
XX
AC   AAC38126;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 19870.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
```

```
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:              2.56e-20        Length:         471
Score:                  247.50          Matches:        48
Percent Similarity:     70.59%          Conservative:   12
Best Local Similarity:  56.47%          Mismatches:     22
Query Match:            46.70%          Indels:         3
DB:                     3               Gaps:           1

US-10-629-953-4 (1-101) x AAC38126 (1-471)

Qy         15 ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAsp 34
              ||||||   |||            ||||||                  ||||||
Db        175 CCGGCTAAGAAAGTCGCAAAGCCAGCAGACGAT---------TCCGACCAGTCTGACGAT 225

Qy         35 IleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLys 54
              ||||||||       ||||||||||||||||||||:::||||||   ||||||||||||
Db        226 ATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAG 285

Qy         55 ValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74
              :::   :::||||||||||||||||||   |||||||||||||||||||:::|||
Db        286 ATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAA 345

Qy         75 GlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAsp 94
              ||||||   |||   :::|||||||||       ||||||::::::   :::   |||:::
Db        346 GGTATCTCTCTAAGCGTGGATCAGTGGAACACTCTTCGGAACCACGCAGAGGATATCGAG 405

Qy         95 GluAlaIleLysGlu 99
              :::|||:::   :::
Db        406 AAGGCCCTCTCTGAC 420

RESULT 10
AAZ50416
ID   AAZ50416 standard; cDNA; 516 BP.
XX
AC   AAZ50416;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Vernonia transcription coactivator PC4(P15) type 1 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 1;
KW   vernonia; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone vs1n.pk013.f21; ss.
XX
OS   Vernonia mespilifolia.
XX
FH   Key             Location/Qualifiers
FT   mat_peptide     1. .333
FT                   /*tag= a
FT                   /product= "Vernonia PC4(P15) type I transcription
FT                   coactivator"
FT                   /note= "Derived from clone vs1n.pk013.f21"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44883.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 3; Page 38; 43pp; English.
XX
CC   The present sequence is the cDNA encoding vernonia Positive Cofactor 4
CC   (PC4) transcription coactivator, designated as PC4(P15) type 1. It is
CC   isolated from clone vs1n.pk013.f21, obtained from vs1n cDNA library,
CC   prepared from vernonia seeds. PC4 molecules functions as a signal
CC   mediator between activators like Gal4/VP16 and general transcription
CC   factors (GTFs) in a transcription initiation complex. The expression
CC   levels of PC4 can be manipulated and the functional properties of
CC   specific transcriptional activators can be modulated. The PC4 DNA can be
CC   used to create transgenic plants with altered PC4 levels, that would
CC   affect the level of transcription of specific genes in the plant. It is
CC   also used for immunological screening of cDNA libraries and to raise
```

```
CC   specific antibodies for detection. The DNA is used as probes and primers,
CC   for genetic and physical mapping of genes and as markers for traits
CC   linked to those genes
XX
SQ   Sequence 516 BP; 153 A; 103 C; 121 G; 126 T; 0 U; 13 Other;

Alignment Scores:
Pred. No.:              2.26e-15       Length:         516
Score:                  206.50         Matches:        40
Percent Similarity:     64.20%         Conservative:   12
Best Local Similarity:  49.38%         Mismatches:     26
Query Match:            38.96%         Indels:         3
DB:                     3              Gaps:           1

US-10-629-953-4 (1-101) x AAZ50416 (1-516)

Qy         17 AlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleVal 36
              :::|||:::      |||||||||||||          |||          |||||||||
Db         97 TCCAAGAAAGACACCGCCGGTGATGATTCCGACGACGAA---------GACGACATCTTC 147

Qy         37 ValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValVal 56
              :::      :::|||||||||||||||||:::||||||   |||    |||:::|||
Db        148 ATTTGTGACGTTTCCAAAAACCGGAGGGTTTCTGTCAGGAANTGGCAAGGGAGGGTCTTT 207

Qy         57 ValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIle 76
              ||||||||||||   ||||||   |||   |||       :::|||   :::   ||||||
Db        208 GTCGATATCCGTGNGTTTTACATGAAANACGGCAANNAAATGCCTGNCAAAANAGGCATC 267

Qy         77 GlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAla 96
                 |||   ||||||   ||||||   ||||||   ::::::          |||:::|||
Db        268 TCATTGACCATGGATCANTGGAAAGAACTCCGTGCTCATGTGGATGAAANTGACAAGGCG 327

Qy         97 Ile 97
              :::
Db        328 CTT 330

RESULT 11
AAC32760
ID   AAC32760 standard; DNA; 1063 BP.
XX
AC   AAC32760;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 548.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
```

PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	1.52e-13	Length:	1063
Score:	194.50	Matches:	39
Percent Similarity:	70.49%	Conservative:	4
Best Local Similarity:	63.93%	Mismatches:	15
Query Match:	36.70%	Indels:	3
DB:	3	Gaps:	1

US-10-629-953-4 (1-101) x AAC32760 (1-1063)

```
Qy         15 ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAsp 34
              ||||||   |||           ||||||                        ||||||
Db        175 CCGGCTAAGAAAGTCGCAAAGCCAGCAGACGAT---------TCCGACCAGTCTGACGAT 225

Qy         35 IleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLys 54
              |||||||||      ||||||||||||||||||||:::||||||   ||||||||||||
Db        226 ATCGTCGTCTGCAACATATCTAAGAATAGGAGAGTCTCTGTAAGGAATTGGAACGGGAAG 285

Qy         55 ValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLys 74
              :::   :::||||||||||||||||||   ||||||||||||||||||||||:::|||
Db        286 ATTTGGATTGACATTCGTGAGTTCTATGTCAAGGACGGAAAGACTTTGCCTGGCAAGAAA 345

Qy         75 Gly 75
              |||
Db        346 GGT 348
```

RESULT 12
ACF03526
ID ACF03526 standard; cDNA; 913 BP.
XX
AC ACF03526;
XX
DT 27-OCT-2003 (revised)
DT 12-SEP-2003 (first entry)
XX
DE Brassica campestris MIP102 encoding cDNA SEQ ID NO:1.
XX
KW Brassica campestris; plant; MIP102; virus resistance;
KW plant virus transport protein binding protein; gene; ss.
XX
OS Brassica rapa.
XX
FH Key   Location/Qualifiers
FT CDS   14. .511
FT   /*tag= a
FT   /product= "MIP102"
XX
PN WO2003022039-A1.
XX
PD 20-MAR-2003.
XX
PF 10-SEP-2001; 2001WO-JP007858.
XX

```
PR   10-SEP-2001; 2001WO-JP007858.
XX
PA   (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PI   Nishiguchi M,  Nyunoya H,  Matsushita Y;
XX
DR   WPI; 2003-313169/30.
DR   P-PSDB; ABR57409.
XX
PT   Plant protein binding to plant virus transport protein for imparting
PT   virus resistance to plants including tobacco.
XX
PS   Claim 3; Fig 2; 84pp; Japanese.
XX
CC   The present invention describes a method for imparting virus resistance
CC   to plants in which the plant cells are transformed with a polynucleotide
CC   encoding a protein (I) which binds to plant virus transport protein. Also
CC   described are plants transformed by (I). The method is useful for
CC   increasing resistance of plants such as tobacco, tomato, petunia, pear,
CC   rice, wheat, barley, maize, soybean, oilseed rape, rose, apple, alfalfa,
CC   melon, lavender, onion, spinach and parsley; to viruses including
CC   tobamovirus, tobravirus, dianthovirus, alfamovirus, bromovirus,
CC   cucumovirus, comovirus, nepovirus, caulimovirus, geminivirus, potivirus
CC   and tospovirus. The present sequence encodes Brassica campestris MIP102,
CC   which is a specifically claimed plant virus transport protein binding
CC   protein from the present invention. (Updated on 27-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 913 BP; 306 A; 161 C; 224 G; 222 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.71e-12        Length:         913
Score:                  185.00          Matches:        37
Percent Similarity:     56.70%          Conservative:   18
Best Local Similarity:  38.14%          Mismatches:     22
Query Match:            34.91%          Indels:         20
DB:                     8               Gaps:           2

US-10-629-953-4 (1-101) x ACF03526 (1-913)

Qy          3 ArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAla 22
              ::::::||||||||||   |||
Db        263 AAGGAGGGAAACAAAGAGTTT---------------------------------------- 283

Qy         23 GlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLys 42
                ||||||              |||   |||:::::::::   ::::::|||
Db        284 ---GACGAT-----------------GACGGCGATCTCATCATTTGCAGGCTGTCGGAT 322

Qy         43 AsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPhe 62
              |||||||||   ::::::   :::   ||||||   :::|||   ||||||||||:::
Db        323 AAGAGGAGAGTGACGATTCAGGAGTTTAGAGGAAAGAGTTTGGTTTCCATCAGAGAGTAT 382

Qy         63 TyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGln 82
              |||:::|||||||||||||   ||||||       |||||||||   |||      :::|||
Db        383 TACAAGAAAGACGGCAAAGAGCTTCCTTCTTCTAAAGGAATAAGCTTAACAGACGAACAA 442

Qy         83 TrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              |||          :::   ||||||   ||||||:::   |||:::|||:::
Db        443 TGGTCAACGTTCAAGAAAAATATTCCAGCTATCGAAGCTGCTGTCAAGAAA 493

RESULT 13
AAZ50419
ID   AAZ50419 standard; cDNA; 740 BP.
XX
AC   AAZ50419;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Soybean transcription coactivator PC4(P15) type 2 cDNA.
XX
KW   Positive cofactor 4; PC4; transcription coactivator; PC4(P15) type 2;
KW   soybean; signal mediator; activator; Gal4/VP16; transgenic plant;
KW   general transcription factor; GTF; transcription initiation complex;
KW   immunological screening; detection; marker; clone ses4d.pk0016.g2; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             64. .489
FT                   /*tag=  a
FT                   /product= "Soybean PC4(P15) type 2 transcription
FT                   coactivator"
FT                   /note= "Derived from clone ses4d.pk0016.g2"
XX
PN   WO200005377-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016479.
XX
PR   22-JUL-1998;   98US-0093687P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Caimi PG,  Odell JT,  Sakai H,  Zhu Q;
XX
DR   WPI; 2000-182701/16.
DR   P-PSDB; AAY44886.
XX
PT   Novel PC4 transcriptional coactivator polynucleotides and polypeptides
PT   used to alter the level of PC4(P15) type 1 and PC4(P15) type 2
PT   polynucleotides and polypeptides.
XX
PS   Claim 21; Page 41-42; 43pp; English.
XX
CC   The present sequence is the cDNA encoding soybean Positive Cofactor 4
CC   (PC4) transcription coactivator, designated as PC4(P15) type 2. It is
CC   isolated from clone ses4d.pk0016.g2, obtained from ses4d cDNA library,
CC   prepared from soybean embryogenic suspension 4 days after subculture. PC4
CC   molecules functions as a signal mediator between activators like
CC   Gal4/VP16 and general transcription factors (GTFs) in a transcription
CC   initiation complex. The expression levels of PC4 can be manipulated and
CC   the functional properties of specific transcriptional activators can be
CC   modulated. The PC4 DNA can be used to create transgenic plants with
CC   altered PC4 levels, that would affect the level of transcription of
CC   specific genes in the plant. It is also used for immunological screening
CC   of cDNA libraries and to raise specific antibodies for detection. The DNA
CC   is used as probes and primers, for genetic and physical mapping of genes
CC   and as markers for traits linked to those genes
XX
SQ   Sequence 740 BP; 249 A; 136 C; 182 G; 173 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.51e-12        Length:         740
Score:                  184.50          Matches:        36
Percent Similarity:     61.29%          Conservative:   21
Best Local Similarity:  38.71%          Mismatches:     25
Query Match:            34.81%          Indels:         11
DB:                     3               Gaps:           2

US-10-629-953-4 (1-101) x AAZ50419 (1-740)

Qy          7 LysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGlyAspAspGly 26
              :::      |||||||||              :::|||   :::         |||:::|||
Db        229 GAAGAACTAGGAGGAGGT------------TCCAAGGGCAAGGAGTACGATGATGAAGGC 276

Qy         27 ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArgVal 46
                                  |||:::::::::   ::::::|||      |||||||||
Db        277 --------------------GATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTG 315

Qy         47 AlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyrGluLysAsp 66
                 ::::::   :::   ||||||   :::|||   ||||||||||:::|||:::||||||
Db        316 ACGATTCAGGATTTCAGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGAT 375

Qy         67 GlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeu 86
              ||||||   ||||||       |||||||||   |||      :::||||||
```

```
Db        376 GGCAAGGAACTTCCTACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAGCCTTT 435

Qy         87 ArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              :::   |||:::   ||||||::::::|||||||||:::
Db        436 AAGAAAAATGTGCCTGCCATAGAAAAAGCCATTAAGAAA 474

RESULT 14
ADE81749/c
ID   ADE81749 standard; cDNA; 496 BP.
XX
AC   ADE81749;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Arabidopsis thaliana expressed polynucleotide seq id 520.
XX
KW   genetically modified organism; transgenic organism; plant;
KW   inhibitor testing; activator testing; modifier testing; fungicide;
KW   insecticide; genetic function; genetic regulation; cellular metabolism;
KW   gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   US2003115639-A1.
XX
PD   19-JUN-2003.
XX
PF   26-JAN-2001; 2001US-00770961.
XX
PR   27-JAN-2000; 2000US-0178466P.
XX
PA   (GORL/) GORLACH J.
PA   (ANYY/) AN Y.
PA   (HAMI/) HAMILTON C M.
PA   (PRIC/) PRICE J L.
PA   (RAIN/) RAINES T M.
PA   (YUYY/) YU Y.
PA   (RAME/) RAMEAKA J G.
PA   (PAGE/) PAGE A.
PA   (MATH/) MATHEW A V.
PA   (LEDF/) LEDFORD B L.
PA   (WOES/) WOESSNER J P.
PA   (HAAS/) HAAS W D.
PA   (GARC/) GARCIA C A.
PA   (KRIC/) KRICKER M.
PA   (SLAT/) SLATER T.
PA   (DAVI/) DAVIS K R.
PA   (ALLE/) ALLEN K.
PA   (HOFF/) HOFFMAN N.
PA   (HURB/) HURBAN P.
XX
PI   Gorlach J,  An Y,  Hamilton CM,  Price JL,  Raines TM,  Yu Y;
PI   Rameaka JG,  Page A,  Mathew AV,  Ledford BL,  Woessner JP,  Haas WD;
PI   Garcia CA,  Kricker M,  Slater T,  Davis KR,  Allen K,  Hoffman N;
PI   Hurban P;
XX
DR   WPI; 2003-810930/76.
XX
PT   Novel Arabidopsis thaliana nucleic acids useful for generating
PT   genetically modified transgenic organisms, for screening biologically
PT   active agents such as fungicides, insecticides.
XX
PS   Claim 1; SEQ ID NO 520; 44pp; English.
XX
CC   The invention describes a nucleic acid (I) comprising a sequence capable
CC   of hybridising under stringent conditions to any one of 999 fully defined
CC   Arabidopsis thaliana sequences (1) as given in specification e.g., 360,
CC   1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a
CC   hybridisation probe to complementary molecules in a cDNA library. (I) is
CC   also useful for generating genetically modified and transgenic organisms,
CC   usually plant cells and plants. A proten encoded by (I) is useful in
CC   screening assays to determine the effect of candidate inhibitors,
CC   activators or modifiers of the gene product. The protein is also useful
CC   for screening biologically active agents e.g., fungicides and
CC   insecticides. A genetically modified cell, comprising an exogenous
CC   nucleic acid, where the nucleic acid comprises transcription regulatory
CC   sequences operably linked to a sequence capable of hybridising under
CC   stringent conditions to (1) is useful in the study of genetic function
CC   and regulation, for alteration of the cellular metabolism and for
CC   screening compounds that may affect the biological function of the gene
CC   or gene product. This sequence represents an Arabidopsis thaliana
CC   polynucleotide of the invention.
XX
SQ   Sequence 496 BP; 147 A; 108 C; 77 G; 164 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.38e-12        Length:          496
Score:                    183.00          Matches:         37
Percent Similarity:       56.25%          Conservative:    17
Best Local Similarity:    38.54%          Mismatches:      22
Query Match:              34.53%          Indels:          20
DB:                       10              Gaps:            2

US-10-629-953-4 (1-101) x ADE81749 (1-496)

Qy          4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
              ||||||||||||   |||
Db        459 AAAGGAAACAAAGAGTTT---------------------------------------- 442

Qy         24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
              ||||||                  |||   |||:::::::::   ::::::|||
Db        441 GATGAT----------------GACGGCGATCTTATCATTTGCAGGTTATCGGATAAG 400

Qy         44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||||||||   ::::::   :::   ||||||   :::|||   ||||||||||:::|||
Db        399 AGAAGAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGAGTATTAC 340

Qy         64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83
              :::||||||||||||   ||||||       |||||||||   |||       :::||||||
Db        339 AAGAAAGATGGCAAAGAACTTCCTACTTCTAAAGGAATAAGCTTAACAGATGAACAATGG 280

Qy         84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                    :::   |||:::   ||||||:::   |||:::|||:::
Db        279 TCAACCTTCAAGAAAAACATGCCAGCCATCGAAAATGCTGTCAAGAAA 232

RESULT 15
AAC50198
ID   AAC50198 standard; DNA; 757 BP.
XX
AC   AAC50198;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 63940.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
```

PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	2.34e-12	Length:	757
Score:	183.00	Matches:	37
Percent Similarity:	56.25%	Conservative:	17
Best Local Similarity:	38.54%	Mismatches:	22
Query Match:	34.53%	Indels:	20
DB:	3	Gaps:	2

US-10-629-953-4 (1-101) x AAC50198 (1-757)

```
Qy          4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
              ||||||||||||   |||
Db        343 AAAGGAAACAAAGAGTTT---------------------------------------- 360

Qy         24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
              ||||||                  |||   |||:::::::::   :::::::|||
Db        361 GATGAT-----------------GACGGCGATCTTATCATTTGCAGGTTATCGGATAAG 402

Qy         44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||||||||   ::::::   :::   ||||||   :::|||   ||||||||||:::|||
Db        403 AGAAGAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGAGTATTAC 462

Qy         64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83
              :::||||||||||||   ||||||       |||||||||   |||       :::||||||
Db        463 AAGAAAGATGGCAAAGAACTTCCTACTTCTAAAGGAATAAGCTTAACAGATGAACAATGG 522

Qy         84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                    :::   |||:::   ||||||:::   |||:::|||:::
Db        523 TCAACCTTCAAGAAAAACATGCCAGCCATCGAAAATGCTGTCAAGAAA 570
```

Search completed: January 12, 2005, 02:45:17
Job time : 417 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2005, 01:26:01 ; Search time 88 Seconds
(without alignments)
815.792 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10629953/runat_10012005_172536_9437/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10629953_@CGN_1_1_69_@runat_10012005_172536_9437 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	368	69.4	289	4	US-09-313-294A-1459	Sequence 1459, Ap
	2	297	56.0	263	4	US-09-313-294A-1612	Sequence 1612, Ap
	3	275	51.9	286	4	US-09-313-294A-6916	Sequence 6916, Ap
	4	160.5	30.3	674	4	US-09-854-133-477	Sequence 477, App
	5	153.5	29.0	662	4	US-09-513-999C-14451	Sequence 14451, A
	6	136.5	25.8	452	4	US-09-621-976-18629	Sequence 18629, A
	7	106	20.0	480	4	US-09-248-796A-5030	Sequence 5030, Ap
	8	84	15.8	325	4	US-09-621-976-18626	Sequence 18626, A
	9	84	15.8	400	4	US-09-513-999C-14452	Sequence 14452, A
	10	84	15.8	479	4	US-09-621-976-18624	Sequence 18624, A
	11	79.5	15.0	1320	4	US-09-724-797-85	Sequence 85, Appl
c	12	75	14.2	2058	4	US-09-252-991A-15140	Sequence 15140, A
	13	75	14.2	2211	4	US-09-252-991A-14623	Sequence 14623, A
c	14	75	14.2	2712	4	US-09-919-172-40	Sequence 40, Appl
c	15	74.5	14.1	528	4	US-09-489-039A-1584	Sequence 1584, Ap
c	16	74.5	14.1	1230	4	US-09-489-039A-1484	Sequence 1484, Ap
	17	74.5	14.1	1737	4	US-09-489-039A-1730	Sequence 1730, Ap
	18	74	14.0	828	4	US-09-543-681A-3276	Sequence 3276, Ap
	19	74	14.0	1524	4	US-09-489-039A-5551	Sequence 5551, Ap
	20	73.5	13.9	1224	4	US-09-266-965-22	Sequence 22, Appl
	21	73.5	13.9	12249	4	US-09-266-965-74	Sequence 74, Appl
	22	73.5	13.9	18331	4	US-09-266-965-96	Sequence 96, Appl
c	23	73.5	13.9	640681	4	US-09-790-988-1	Sequence 1, Appli
	24	73.5	13.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	25	73.5	13.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	26	72.5	13.7	4043	4	US-09-814-915A-70	Sequence 70, Appl
	27	72	13.6	226	4	US-09-621-976-13408	Sequence 13408, A
	28	72	13.6	547	4	US-09-513-999C-148	Sequence 148, App
	29	72	13.6	649	4	US-09-513-999C-147	Sequence 147, App
	30	72	13.6	714	4	US-09-252-991A-674	Sequence 674, App
c	31	72	13.6	951	4	US-09-252-991A-741	Sequence 741, App
c	32	71.5	13.5	16826	4	US-08-956-171E-46	Sequence 46, Appl
c	33	71.5	13.5	16826	4	US-08-781-986A-46	Sequence 46, Appl
c	34	71	13.4	474	4	US-09-252-991A-15005	Sequence 15005, A
	35	70.5	13.3	837	4	US-09-252-991A-15872	Sequence 15872, A
	36	70.5	13.3	972	4	US-09-248-796A-1330	Sequence 1330, Ap
c	37	70.5	13.3	1158	4	US-09-252-991A-15820	Sequence 15820, A
	38	70.5	13.3	2366	4	US-09-799-451-351	Sequence 351, App
	39	70.5	13.3	2970	4	US-09-252-991A-15898	Sequence 15898, A
	40	70	13.2	5173	3	US-08-801-308-2	Sequence 2, Appli
c	41	70	13.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	42	70	13.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	43	69.5	13.1	162450	3	US-09-345-882-1	Sequence 1, Appli
c	44	69	13.0	10709	4	US-09-596-002-9	Sequence 9, Appli
	45	68.5	12.9	43280	2	US-08-804-227C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-313-294A-1459
; Sequence 1459, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1459
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550983H1
US-09-313-294A-1459

Alignment Scores:
Pred. No.: 1.05e-36 Length: 289
Score: 368.00 Matches: 70
Percent Similarity: 86.17% Conservative: 11
Best Local Similarity: 74.47% Mismatches: 11
Query Match: 69.43% Indels: 2
DB: 4 Gaps: 1

US-10-629-953-4 (1-101) x US-09-313-294A-1459 (1-289)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||  ||||||   ||||||||||||||||||       ||||||||||||||||||
Db          6 ATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 65
```

```
Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   |||||||||||||||||||||||:::|||      |||||||||
Db         66 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 125

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||||:::::::|||:::||||||:::|||   |||:::|||   ||||||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGAC 185

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||   |||||||||||||||||||||   |||||||||||   |||
Db        186 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 245

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAla 92
                 :::|||||||||||||||||:::|||||||||||||||
Db        246 CAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCC 287
```

RESULT 2
US-09-313-294A-1612
; Sequence 1612, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1612
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551220H1
; NAME/KEY: unsure
; LOCATION: 25, 181
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-1612

Alignment Scores:

Pred. No.:	4.72e-28	Length:	263
Score:	297.00	Matches:	60
Percent Similarity:	82.56%	Conservative:	11
Best Local Similarity:	69.77%	Mismatches:	13
Query Match:	56.04%	Indels:	3
DB:	4	Gaps:	1

US-10-629-953-4 (1-101) x US-09-313-294A-1612 (1-263)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||  ||||||      ||||||||||||||      ||||||||||||||||||
Db          6 ATGTGGGGGAAAGGAAAGANGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 65

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   |||||||||||||||||||||||:::|||      |||||||||
Db         66 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 125

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||||:::::::|||:::||||||:::|||   |||:::|||   |||||
Db        126 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGNCGAC 185

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||||   |||||||||||||||||||||   |||||||||||   :::
Db        186 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATC-ATC 244

Qy         79 ProMetAspGlnTrpLys 84
                 :::|||||||||||||
Db        245 CAATTAGATCAGTGGAAG 262
```

RESULT 3
US-09-313-294A-6916
; Sequence 6916, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6916
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700380835H1
; NAME/KEY: unsure
; LOCATION: 54
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6916

Alignment Scores:

Pred. No.:	2.63e-25	Length:	286
Score:	275.00	Matches:	57
Percent Similarity:	83.54%	Conservative:	9
Best Local Similarity:	72.15%	Mismatches:	10
Query Match:	51.89%	Indels:	3
DB:	4	Gaps:	1

US-10-629-953-4 (1-101) x US-09-313-294A-6916 (1-286)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPhe-GlyGlyGly------GlyGluProAlaAlaLy 18
              |||     ||||||   |||||||| ||||||||      ||||||||||||||||||
Db         50 ATGTNGGGGAAAGGAAAGAAGCGTTTCGGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAA 109

Qy         18 sArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAl 38
              ||||:::||||||   |||||||||||||||||||||:::|||      ||||||||
Db        110 GCGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGC 169

Qy         38 aGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAs 58
              |:::|||||||||||:::::::|||:::||||||:::|||   |||:::|||   |||||
Db        170 CGAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGA 229

Qy         58 pIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIle 76
              |:::|||||||||||   |||||||||||||||||||   |||||||||||
Db        230 CTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATA 284
```

RESULT 4
US-09-854-133-477
; Sequence 477, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 477
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Homo sapiens

```
US-09-854-133-477

Alignment Scores:
Pred. No.:                9.01e-11        Length:         674
Score:                    160.50          Matches:        36
Percent Similarity:       60.00%          Conservative:   21
Best Local Similarity:    37.89%          Mismatches:     27
Query Match:              30.28%          Indels:         11
DB:                       4               Gaps:           3

US-10-629-953-4 (1-101) x US-09-854-133-477 (1-674)

Qy          14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGly---------------------- 26
               |||          |||:::::::      |||:::
Db         148 GAAAAACCTGTAAAGAAACAAAAGACAGGTGAGACTTCGAGAGCCCTGTCATCTTCTAAA 207

Qy          27 ---ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
                  |||    |||   ||||||::::::          ||||||   |||   |||
Db         208 CAGAGCAGCAGCAGCAGAGATGATAACATGTTT------CAGATTGGGAAAATGAGGTAC 261

Qy          46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr---Glu 64
               |||:::||||||   :::   |||||||||::::::||||||||||||::::::   :::
Db         262 GTTAGTGTTCGCGATTTTAAAGGCAAAGTGCTAATTGATATTAGAGAATATTGGATGGAT 321

Qy          65 LysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLys 84
                  :::|||:::      ||||||||||||||||||   |||       :::||||||
Db         322 CCTGAAGGTGAAATGAAACCAGGAAGAAAAGGTATTTCTTTAAATCCAGAACAATGGAGC 381

Qy          85 IleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                  |||::::::   |||       ||||||:::|||:::::::::
Db         382 CAGCTGAAGGAACAGATTTCTGACATTGATGATGCAGTAAGAAAA 426

RESULT 5
US-09-513-999C-14451
; Sequence 14451, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14451
;   LENGTH: 662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 431
;   OTHER INFORMATION: m=a or c
US-09-513-999C-14451

Alignment Scores:
Pred. No.:                6.35e-10        Length:         662
Score:                    153.50          Matches:        35
Percent Similarity:       58.95%          Conservative:   21
Best Local Similarity:    36.84%          Mismatches:     28
Query Match:              28.96%          Indels:         11
DB:                       4               Gaps:           3

US-10-629-953-4 (1-101) x US-09-513-999C-14451 (1-662)

Qy          14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGly---------------------- 26
               |||          |||:::::::      |||:::
Db         169 GAAAAACCTGTAAAGAAACAAAAGACAGGTGAGACTTCGAGAGCCCTGTCATCTTCTAAA 228

Qy          27 ---ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
                  |||    |||   ||||||::::::          ||||||   |||   |||
Db         229 CAGAGCAGCAGCAGCAGAGATGATAACATGTTT------CAGATTGGGAAAATGAGGTAC 282

Qy          46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr---Glu 64
               |||:::||||||   :::   |||||||||::::::||||||||||||::::::   :::
Db         283 GTTAGTGTTCGCGATTTTAAAGGCAAAGTGCTAATTGATATTAGAGAATATTGGATGGAT 342

Qy          65 LysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLys 84
                  :::|||:::      ||||||||||||||||||   |||       :::||||||
Db         343 CCTGAAGGTGAAATGAAACCAGGAAGAAAAGGTATTTCTTTAAATCCAGAACAATGGAGC 402

Qy          85 IleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                  |||::::::   |||       |||   :::|||:::::::::
Db         403 CAGCTGAAGGAACAGATTTCTGACATTGMTGATGCAGTAAGAAAA 447

RESULT 6
US-09-621-976-18629
; Sequence 18629, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 18629
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-18629

Alignment Scores:
Pred. No.:                4.66e-08        Length:         452
Score:                    136.50          Matches:        36
Percent Similarity:       58.95%          Conservative:   20
Best Local Similarity:    37.89%          Mismatches:     28
Query Match:              25.75%          Indels:         12
DB:                       4               Gaps:           3

US-10-629-953-4 (1-101) x US-09-621-976-18629 (1-452)

Qy          14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGly---------------------- 26
               |||          |||::::::      |||:::
Db         137 GAAAAACCTGTAAAGAAACAAAAGACAGGTGAGACTTCGAGAGCCCTGTCATCTTCTAAA 196

Qy          27 ---ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
                  |||    |||   ||||||::::::          ||||||   |||   |||
Db         197 CAGAGCAGCAGCAGCAGAGATGATAACATGTTT------CAGATTGGGAAAATGAGGTAC 250

Qy          46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr---Glu 64
               |||   ||||||   :::   |||||||||::::::||||||||||||::::::   :::
Db         251 GTTAT-GTTCGCGATTTTAAAGGCAAAGTGCTAATTGATATTAGAGAATATTGGATGGAT 309

Qy          65 LysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLys 84
                  :::|||:::      ||||||||||||||||||   |||       :::||||||
Db         310 CCTGAAGGTGAAATGAAACCAGGAAGAAAAGGTATTTCTTTAAATCCAGAACAATGGAGC 369

Qy          85 IleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                  |||::::::   |||       ||||||:::|||:::::::::
Db         370 CAGCTGAAGGAACAGATTTCTGACATTGATGATGCAGTAAGAAAA 414

RESULT 7
US-09-248-796A-5030
; Sequence 5030, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
```

```
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5030
;   LENGTH: 480
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-5030

Alignment Scores:
Pred. No.:              0.000277        Length:         480
Score:                  106.00          Matches:        29
Percent Similarity:     55.29%          Conservative:   18
Best Local Similarity:  34.12%          Mismatches:     30
Query Match:            20.00%          Indels:         8
DB:                     4               Gaps:           3

US-10-629-953-4 (1-101) x US-09-248-796A-5030 (1-480)

Qy         16 AlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspIle 35
              |||   |||||   |||           |||:::   ::::::::::::
Db         43 GCATTCAAAAGAGGAGCA--------------CCAGCATCCAATTCAAATGAAGTAGAG 87

Qy         36 ValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysVal 55
              ::::::      :::   |||      :::|||   ||||||   :::|||
Db         88 ATTTTGTTGGATCTGAAGAAA------CAGGTTACTGTTCGTAAATACAACAACGTCAAT 141

Qy         56 ValValAspIleArgGluPheTyr---GluLysAspGlyLysThrLeuProGlyArgLys 74
              :::||||||||||||||||||||   ::::::||||||         ||||||:::|||
Db        142 TTAGTAGATATTAGAGAGTTTTATACCGATAGGGATGGCGTCAAGAGACCAGGTAAAAAG 201

Qy         75 GlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAsp 94
              ||||||   |||       |||   :::       |||   :::             |||
Db        202 GGAATATCATTGACTGAAGACACTTACTATAAGTTGTTGGAAGCTACTAACAAGATTCAA 261

Qy         95 GluAlaIleLysGlu 99
              |||:::   :::
Db        262 AACGCCTTAGACGAT 276

RESULT 8
US-09-621-976-18626
; Sequence 18626, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 18626
;   LENGTH: 325
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-18626

Alignment Scores:
Pred. No.:              0.0825          Length:         325
Score:                  84.00           Matches:        21
Percent Similarity:     55.17%          Conservative:   11
Best Local Similarity:  36.21%          Mismatches:     16
Query Match:            15.85%          Indels:         10
DB:                     4               Gaps:           2

US-10-629-953-4 (1-101) x US-09-621-976-18626 (1-325)

Qy         14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGly---------------------- 26
              |||         |||::::::      |||:::
Db        137 GAAAAACCTGTAAAGAAACAAAAGACAGGTGAGACTTCGAGAGCCCTGTCATCTTCTAAA 196

Qy         27 ---ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
                  |||   |||   ||||||::::::         ||||||   |||   |||
Db        197 CAGAGCAGCAGCAGCAGAGATGATAACATGTTT------CAGATTGGGAAAATGAGGTAC 250

Qy         46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||:::||||||   :::   |||||||||::::::||||||||||:::::
Db        251 GTTAGTGTTCGCGATTTTAAAGGCAAAGTGCTAATTGATATTAGAGAATATTGG 304

RESULT 9
US-09-513-999C-14452
; Sequence 14452, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14452
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 71
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 72
;   OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-14452

Alignment Scores:
Pred. No.:              0.109           Length:         400
Score:                  84.00           Matches:        21
Percent Similarity:     55.17%          Conservative:   11
Best Local Similarity:  36.21%          Mismatches:     16
Query Match:            15.85%          Indels:         10
DB:                     4               Gaps:           2

US-10-629-953-4 (1-101) x US-09-513-999C-14452 (1-400)

Qy         14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGly---------------------- 26
              |||         |||::::::      |||:::
Db        171 GAAAAACCTGTAAAGAAACAAAAGACAGGTGAGACTTCGAGAGCCCTGTCATCTTCTAAA 230

Qy         27 ---ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
                  |||   |||   ||||||::::::         ||||||   |||   |||
Db        231 CAGAGCAGCAGCAGCAGAGATGATAACATGTTT------CAGATTGGGAAAATGAGGTAC 284

Qy         46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||:::||||||   :::   |||||||||::::::||||||||||:::::
Db        285 GTTAGTGTTCGCGATTTTAAAGGCAAAGTGCTAATTGATATTAGAGAATATTGG 338

RESULT 10
US-09-621-976-18624
; Sequence 18624, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
```

```
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 18624
;   LENGTH: 479
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-18624

Alignment Scores:
Pred. No.:              0.138          Length:        479
Score:                  84.00          Matches:       21
Percent Similarity:     55.17%         Conservative:  11
Best Local Similarity:  36.21%         Mismatches:    16
Query Match:            15.85%         Indels:        10
DB:                     4              Gaps:          2

US-10-629-953-4 (1-101) x US-09-621-976-18624 (1-479)

Qy         14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGly---------------------- 26
              |||         |||::::::      |||:::
Db        167 GAAAAACCTGTAAAGAAACAAAAGACAGGTGAGACTTCGAGAGCCCTGTCATCTTCTAAA 226

Qy         27 ---ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
                 |||   |||   ||||||::::::         ||||||   |||   |||
Db        227 CAGAGCAGCAGCAGCAGAGATGATAACATGTTT------CAGATTGGGAAAATGAGGTAC 280

Qy         46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||:::||||||   :::   |||||||||:::::::||||||||||||:::::
Db        281 GTTAGTGTTCGCGATTTTAAAGGCAAAGTGCTAATTGATATTAGAGAATATTGG 334

RESULT 11
US-09-724-797-85
; Sequence 85, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
;  APPLICANT: Jon S. THORSON
;  TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
;  TITLE OF INVENTION:  ENCODING FOR BIOSYNTHESIS OF
;  TITLE OF INVENTION:  CALICHEAMICIN AND SELF-RESISTANCE THERETO
;  FILE REFERENCE: 2653-40
;  CURRENT APPLICATION NUMBER: US/09/724,797
;  CURRENT FILING DATE:  2000-11-28
;  PRIOR APPLICATION NUMBER: 60/111,325
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 85
;   LENGTH: 1320
;   TYPE: DNA
;   ORGANISM: .
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1320)
US-09-724-797-85

Alignment Scores:
Pred. No.:              1.88           Length:        1320
Score:                  79.50          Matches:       26
Percent Similarity:     43.68%         Conservative:  12
Best Local Similarity:  29.89%         Mismatches:    44
Query Match:            15.00%         Indels:        5
DB:                     4              Gaps:          2

US-10-629-953-4 (1-101) x US-09-724-797-85 (1-1320)

Qy          4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
              :::|||      |||   |||      |||      |||         :::      |||

Db        957 CGCGGCGCGGGGCGGGGTGGGCGGCCGGGACGCGGCGCAGGGGCGGAGCAAGGTCCGGGG 1016

Qy         24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
                 |||          :::|||          ::::::||||||
Db       1017 CCGCGTGGCGCTGGCCACGCTGCTGGCGGCCGGCTGCTGGTGGCAGCCGCCGGTGCCCCC 1076

Qy         44 ArgArgValAlaValArg---------ThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||   |||   |||                  |||:::|||:::      :::|||
Db       1077 CGCCGACCTGCGCAGCGTCGACGACCCGACCGGCCGGGCCGCGTACTTGCGCCGCTGCGG 1136

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||         ||||||:::   |||   ||||||      :::
Db       1137 GAGTTCCTC------GACGGGCAGCGGCTCACCGGCCGGGTCGAGGTGCCGCCGACCCGC 1190

Qy         81 AspGlnTrpLysIleLeuArg 87
              :::   |||:::      |||
Db       1191 AACTACTGGGAGGCGGCGCGG 1211

RESULT 12
US-09-252-991A-15140/c
; Sequence 15140, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 15140
;   LENGTH: 2058
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15140

Alignment Scores:
Pred. No.:              12.1           Length:        2058
Score:                  75.00          Matches:       26
Percent Similarity:     46.67%         Conservative:  9
Best Local Similarity:  34.67%         Mismatches:    20
Query Match:            14.15%         Indels:        20
DB:                     4              Gaps:          4

US-10-629-953-4 (1-101) x US-09-252-991A-15140 (1-2058)

Qy          2 TrpArgLysGlyAsnLysArgPheGlyGly---------GlyGlyGlu------------ 14
              :::|||   |||      ||||||||||||         |||||||||
Db        252 TTCCGCGGCGGGCGCCTGCGCTTCGGCGGGAGCGGCGCTGGCGGGGAGCGCGACGCCGAG 193

Qy         15 ------ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAsp 32
                   |||:::|||||||||||||||   |||      :::      |||
Db        192 GAAGGCGAGGGCGCCGAGCGCCGCGCCGCCGGCCAGGACCTGGCGGCGGCCGAGGCCGGC 133

Qy         33 AspAspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsn 52
              |||   :::|||      ::::::                        ||||||
Db        132 GATCAGGTCGTTCATGTGCGGGTT--------------------GTTCGAATGGTTGCT 94

Qy         53 GlyLysValValValAspIleArgGluPheTyrGluLysAspGly 67
              |||::::::|||                  |||:::||||||
Db         93 GGGCAACTCGTCGCC------------ATCACCGAACAGGATGGC 61

RESULT 13
US-09-252-991A-14623
; Sequence 14623, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

```
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14623
;   LENGTH: 2211
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14623

Alignment Scores:
Pred. No.:                13.3          Length:          2211
Score:                    75.00         Matches:         26
Percent Similarity:       46.67%        Conservative:    9
Best Local Similarity:    34.67%        Mismatches:      20
Query Match:              14.15%        Indels:          20
DB:                       4             Gaps:            4

US-10-629-953-4 (1-101) x US-09-252-991A-14623 (1-2211)

Qy          2 TrpArgLysGlyAsnLysArgPheGlyGly---------GlyGlyGlu------------ 14
              :::|||   |||       ||||||||||||         |||||||||
Db       1906 TTCCGCGGCGGGCGCCTGCGCTTCGGCGGGAGCGGCGCTGGCGGGGAGCGCGACGCCGAG 1965

Qy         15 ------ProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAsp 32
                    |||:::||||||||||||||   |||     :::      |||
Db       1966 GAAGGCGAGGGCGCCGAGCGCCGCGCCGCCGGCCAGGACCTGGCGGCGGCCGAGGCCGGC 2025

Qy         33 AspAspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsn 52
              |||   :::|||      ::::::                        ||||||
Db       2026 GATCAGGTCGTTCATGTGCGGGTT--------------------GTTCGAATGGTTGCT 2064

Qy         53 GlyLysValValValAspIleArgGluPheTyrGluLysAspGly 67
              |||::::::|||                    |||:::||||||
Db       2065 GGGCAACTCGTCGCC------------ATCACCGAACAGGATGGC 2097

RESULT 14
US-09-919-172-40/c
; Sequence 40, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
;  APPLICANT: Faris, Mary
;  APPLICANT:  Turner, Christopher M.
;  TITLE OF INVENTION: PROSTATE CANCER MARKERS
;  FILE REFERENCE: PA-0036 US
;  CURRENT APPLICATION NUMBER: US/09/919,172
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: 60/222,469
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PERL Program
; SEQ ID NO 40
;   LENGTH: 2712
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673545 1798379CB1
US-09-919-172-40

Alignment Scores:
Pred. No.:                17.4          Length:          2712
Score:                    75.00         Matches:         31
Percent Similarity:       37.19%        Conservative:    14
Best Local Similarity:    25.62%        Mismatches:      37
Query Match:              14.15%        Indels:          39
DB:                       4             Gaps:            5

US-10-629-953-4 (1-101) x US-09-919-172-40 (1-2712)

Qy          1 MetTrpArgLysGlyAsn---LysArgPheGlyGlyGlyGly----------------- 13
              :::|||:::   ||||||   ||||||   ||||||||||
Db       2292 ATCTGGAAAATGGGCAATGCAAAGAGGCATGGTGGGGGCGGGAATGGGCATGGCTAGAAG 2233

Qy         14 ----------------------GluProAlaAlaLysArgArg-----------Ala 21
                                    |||      :::||||||
Db       2232 CACAAACCTAGACAGCACAGATTCCGCCCCAGAACACAGAGGCGTCACTGGCTCTAACTA 2173

Qy         22 AlaGlyAspAspGly---------------------------------------------- 26
                |||   :::
Db       2172 ACTGGTTTGGAGAGAAGATCTTTCAAAAACTGTATGTGGACATTCTTTTCCCTCAAAGTG 2113

Qy         27 ProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArgVal 46
              |||||||||     |||   ::::::        |||:::          |||:::
Db       2112 CCATCAGAGATAAATGACAAAAATTTAACTTGCTGCCAGTTAGAGGTACAGAGACAGGCT 2053

Qy         47 AlaVal-----------ArgThrTrpAsnGlyLysVal-ValValAspIleArgGluPh 62
              :::                |||   |||   |||:::||| ||||||
Db       2052 TCTACATTCTGGAGTGGGAGAGTCTGGGTTGGGAGAGTCGGTGGTGGTGGGGGCTCATCT 1993

Qy         62 eTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGl 82
               |||   |||:::      ||||||      :::         :::||||||      ::
Db       1992 TTATCTAAAAGAGAAGGTCACGCTGGAACCAAAAGCTTCACCTGAGCTACCCTTTACAGA 1933

Qy         82 n 82
              :
Db       1932 G 1932

RESULT 15
US-09-489-039A-1584/c
; Sequence 1584, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 1584
;   LENGTH: 528
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1584

Alignment Scores:
Pred. No.:                2.29          Length:          528
Score:                    74.50         Matches:         28
Percent Similarity:       37.76%        Conservative:    9
Best Local Similarity:    28.57%        Mismatches:      34
Query Match:              14.06%        Indels:          27
DB:                       4             Gaps:            5

US-10-629-953-4 (1-101) x US-09-489-039A-1584 (1-528)

Qy          5 GlyAsnLysArgPheGlyGlyGlyGly---------GluProAlaAlaLysArgArgAla 21
              |||::::::          |||||||||          ||||||   |||||||||
Db        399 GGCCACCGGGCCGCCGACGGTGGCGGCCGTGGCCAGCACCCAGCCGGGAAGCGTCGCGGT 340

Qy         22 AlaGlyAspAspGlyProSerGluSer-------------------------AlaAsp 32
              ||||||      ||||||   :::
Db        339 GCGGGTCAGTCGGGTCCAGACGACGGCGAACATCAGCCACAGGGCCACCATCCACAGGGG 280

Qy         33 AspAspIleValValAlaGlnIle------SerLysAsnArgArgValAlaValArgThr 50
                    |||   |||:::|||      |||:::   |||          |||:::|||
Db        279 CAGCAGGCTGTCGCCGCGAAAATCGATCAGTCCCGCCAGCGCCCAGCAGGCGTCCAGACC 220

Qy         51 TrpAsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeu 70
```

```
                                    |||         ::::::   |||:::|||
Db        219 --------------------GCAGCCCGCCGCGGCCAACAGCAGGGCGGGCGGACGACT 181

Qy         71 ------------------ProGlyArgLysGlyIleGlnLeuProMetAspGln 82
                                |||||      |||        |||      |||
Db        180 GGCCGCCGGCAACCGCAGCCAGGCGAAAATGGCCAGCATCAGCCAGATCAGCAG 127
```

Search completed: January 12, 2005, 03:29:36
Job time : 101 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2005, 02:45:26 ; Search time 462 Seconds
(without alignments)
1254.369 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4293498 seqs, 2868903791 residues

Total number of hits satisfying chosen parameters: 8586996

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10629953/runat_10012005_172538_9528/app_query.fasta_1.263
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10629953_@CGN_1_1_480_@runat_10012005_172538_9528
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	530	100.0	1754	17	US-10-437-963-83372	Sequence 83372, A
2	398	75.1	660	16	US-10-425-114-4500	Sequence 4500, Ap
3	395	74.5	757	18	US-10-425-115-74342	Sequence 74342, A
4	304.5	57.5	310	9	US-09-294-093B-67	Sequence 67, Appl
5	269	50.8	663	16	US-10-424-599-38767	Sequence 38767, A
6	261.5	49.3	917	16	US-10-424-599-38768	Sequence 38768, A
c 7	183	34.5	496	10	US-09-770-961-520	Sequence 520, App
8	182.5	34.4	794	16	US-10-425-114-20666	Sequence 20666, A
9	182.5	34.4	911	16	US-10-424-599-4356	Sequence 4356, Ap
c 10	182	34.3	1030	17	US-10-437-963-59601	Sequence 59601, A
11	179	33.8	1479	16	US-10-424-599-9157	Sequence 9157, Ap
c 12	177	33.4	497	16	US-10-424-599-127795	Sequence 127795,
13	172	32.5	2776	13	US-10-087-192-1784	Sequence 1784, Ap
14	166.5	31.4	933	16	US-10-425-114-24336	Sequence 24336, A
15	166.5	31.4	935	16	US-10-425-114-33158	Sequence 33158, A
16	166.5	31.4	1225	18	US-10-425-115-154421	Sequence 154421,
17	166	31.3	658	16	US-10-425-114-21115	Sequence 21115, A
18	166	31.3	658	18	US-10-425-115-20439	Sequence 20439, A
19	163	30.8	483	10	US-09-918-995-25317	Sequence 25317, A
20	160.5	30.3	384	15	US-10-177-293-349	Sequence 349, App
21	160.5	30.3	384	15	US-10-257-021-21	Sequence 21, Appl
22	160.5	30.3	384	18	US-10-370-715B-281	Sequence 281, App
23	160.5	30.3	460	10	US-09-918-995-35177	Sequence 35177, A
24	160.5	30.3	466	13	US-10-087-192-1787	Sequence 1787, Ap
25	160.5	30.3	469	16	US-10-242-535A-39519	Sequence 39519, A
26	160.5	30.3	469	16	US-10-085-783A-39519	Sequence 39519, A
27	160.5	30.3	500	16	US-10-242-535A-51406	Sequence 51406, A
28	160.5	30.3	500	16	US-10-085-783A-51406	Sequence 51406, A
29	160.5	30.3	592	9	US-09-895-828-179	Sequence 179, App
30	160.5	30.3	592	15	US-10-114-666-179	Sequence 179, App
31	160.5	30.3	601	9	US-09-895-828-342	Sequence 342, App
32	160.5	30.3	601	15	US-10-114-666-342	Sequence 342, App
33	160.5	30.3	674	9	US-09-738-973-477	Sequence 477, App
34	160.5	30.3	674	9	US-09-854-133-477	Sequence 477, App
35	160.5	30.3	674	15	US-10-144-649A-477	Sequence 477, App
36	160.5	30.3	685	16	US-10-242-535A-45670	Sequence 45670, A
37	160.5	30.3	685	16	US-10-085-783A-45670	Sequence 45670, A
38	160.5	30.3	689	9	US-09-920-345-29	Sequence 29, Appl
39	160.5	30.3	1387	14	US-10-198-846-13428	Sequence 13428, A
40	160.5	30.3	1391	18	US-10-357-930-25730	Sequence 25730, A
41	160.5	30.3	3527	9	US-09-925-297-320	Sequence 320, App
42	159.5	30.1	400	16	US-10-242-535A-51580	Sequence 51580, A
43	159.5	30.1	400	16	US-10-085-783A-51580	Sequence 51580, A
44	157.5	29.7	432	10	US-09-918-995-6997	Sequence 6997, Ap
45	150.5	28.4	420	16	US-10-242-535A-52202	Sequence 52202, A

ALIGNMENTS

RESULT 1
US-10-437-963-83372
; Sequence 83372, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83372
; LENGTH: 1754

```
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_8270C.1
US-10-437-963-83372

Alignment Scores:
Pred. No.:              3.36e-57        Length:         1754
Score:                  530.00          Matches:        101
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     17              Gaps:           0

US-10-629-953-4 (1-101) x US-10-437-963-83372 (1-1754)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 ATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 1182

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 1242

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 1302

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 GAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATG 1362

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 1422

Qy        101 Ala 101
              |||
Db       1423 GCG 1425

RESULT 2
US-10-425-114-4500
; Sequence 4500, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4500
;   LENGTH: 660
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700380835_FLI
US-10-425-114-4500

Alignment Scores:
Pred. No.:              5.65e-41        Length:         660
Score:                  398.00          Matches:        76
Percent Similarity:     87.25%          Conservative:   13
Best Local Similarity:  74.51%          Mismatches:     11
Query Match:            75.09%          Indels:         2
DB:                     16              Gaps:           1

US-10-629-953-4 (1-101) x US-10-425-114-4500 (1-660)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||    ||||||    |||||||||||||||||      ||||||||||||||||||
Db         51 ATGTGGGGGAAAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 110

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   ||||||||||||||||||||||:::|||      |||||||||
Db        111 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 170

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||:::::::|||:::||||||:::|||   |||:::|||   ||||||
Db        171 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGAC 230

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||||   ||||||||||||||||||||   ||||||||||||   |||
Db        231 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 290

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98
                 :::|||||||||||||||||:::|||||||||||||||||||:::||||||||||:::
Db        291 CAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCCATAAATGAGGCAATCGAG 350

Qy         99 GluAsn 100
              ||||||
Db        351 GAAAAT 356

RESULT 3
US-10-425-115-74342
; Sequence 74342, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 74342
;   LENGTH: 757
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_167799C.1
US-10-425-115-74342

Alignment Scores:
Pred. No.:              1.59e-40        Length:         757
Score:                  395.00          Matches:        75
Percent Similarity:     87.25%          Conservative:   14
Best Local Similarity:  73.53%          Mismatches:     11
Query Match:            74.53%          Indels:         2
DB:                     18              Gaps:           1

US-10-629-953-4 (1-101) x US-10-425-115-74342 (1-757)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              ||||||    :::|||   |||||||||||||||||      ||||||||||||||||||
Db         72 ATGTGGGGGAGAGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGTGAGCCCGCGGCCAAG 131

Qy         19 ArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAla 38
              |||:::||||||   ||||||||||||||||||||||:::|||      |||||||||
Db        132 CGCCAGGCCGCGAGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGCC 191

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||:::::::|||:::||||||:::|||   |||:::|||   ||||||
Db        192 GAGATATCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGAC 251

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              :::|||||||||||||   ||||||||||||||||||||   ||||||||||||   |||
```

```
Db        252 TTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTC 311

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98
              :::||||||||||||||||||:::||||||||||||||||||:::|||||||||:::
Db        312 CAATTAGATCAGTGGAAGATATTGAAGGACAACATCAAAGCCATAAATGAGGCCATCGAG 371

Qy         99 GluAsn 100
              ||||||
Db        372 GAAAAT 377

RESULT 4
US-09-294-093B-67
; Sequence 67, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
;  APPLICANT:      Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:       PERL Program
; SEQ ID NO 67
;   LENGTH: 310
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700281907H2
;   NAME/KEY: unsure
;   LOCATION: 2, 30, 295
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-67

Alignment Scores:
Pred. No.:              1.64e-29        Length:         310
Score:                  304.50          Matches:        65
Percent Similarity:     81.44%          Conservative:   14
Best Local Similarity:  67.01%          Mismatches:     13
Query Match:            57.45%          Indels:         5
DB:                     9               Gaps:           2

US-10-629-953-4 (1-101) x US-09-294-093B-67 (1-310)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGly------GlyGluProAlaAlaLys 18
              |||      :::|||   ||||||||||||||||      ||||||||||||||||||
Db         26 ATGTNGGGGAGGGGAAAGAAGCGTTTCGGCGGTGGCGGCAGCGGCGAGCCCGCGGCCAAG 85

Qy         19 ArgArgAlaAla-GlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAl 38
              |||:::|||||| ||||||||||||||||||||||||||:::|||      ||||||||
Db         86 CGCCAGGCCGCGAGGGGACGATGGCCCCTCCGAATCCGCTGAAGACGGTACCGTCGTAGC 145

Qy         38 aGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAs 58
              |:::|||||||||||:::::|||:::||||||:::|||   |||:::|||   |||||
Db        146 CGAGATATCGAAGAACAAGAAGGTATCCGTTAGGAGCTGGAAAGGCAGGGTCTTCGTCGA 205

Qy         58 pIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGln-L 78
              |:::||||||||||||   ||||||||||||||||||||   ||||||   :::    :
Db        206 CTTACGCGAGTTCTACTTCAAGGACGGCAAGACTCTCCCCACCCGCAAA---GTATATCA 262

Qy         78 euProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIle 93
              ::|||:::      ||||||||||||:::|||   ||||||||||||
Db        263 TTCCAATTAGATCATGGAAGATATTGAAGGACNACATCAAAGCCATA 309

RESULT 5
US-10-424-599-38767
; Sequence 38767, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 38767
;   LENGTH: 663
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_135006C.1
US-10-424-599-38767

Alignment Scores:
Pred. No.:              1.24e-24        Length:         663
Score:                  269.00          Matches:        56
Percent Similarity:     66.99%          Conservative:   13
Best Local Similarity:  54.37%          Mismatches:     14
Query Match:            50.75%          Indels:         20
DB:                     16              Gaps:           3

US-10-629-953-4 (1-101) x US-10-424-599-38767 (1-663)

Qy         17 AlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSer----------------- 30
              ||||||||||||         |||||||||   |||:::::::
Db        178 GCGAAGAGGAGA------GATGACGACGGTGCTTCCGACGCCGACTCTGAAGGCCACGCG 231

Qy         31 --------------------------AlaAspAspAsp---------IleValValAla 38
                                         :::|||||||||         |||   |||
Db        232 CCGCCCAAGAAGTCCCTCAAGAAGGATTCCGATGACGATCCCGACTCTATTACCGTTTGC 291

Qy         39 GlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAsp 58
              :::|||||||||||||||||||||||||||||   |||   |||   ::::::||||||
Db        292 GAGATTTCGAAGAACAGGAGGGTTGCCGTGAGGAACTGGAAAGGCAGCATTATGGTTGAC 351

Qy         59 IleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeu 78
              ||||||||||||||   ||||||||||||   ||||||||||||||||||||||   |||
Db        352 ATTCGCGAGTTTTACGTCAAAGATGGCAAGCAATTGCCTGGCAGAAAAGGTATCTCTTTG 411

Qy         79 ProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLys 98
                 ||||||||||||   :::||||||:::::::::::   ||||||:::||||||
Db        412 ACCATGGATCAGTGGAATGTGCTTCGTAATCATGTTGAAGAAATTGACAAGGCAATTAAT 471

Qy         99 GluAsnAla 101
              ||||||:::
Db        472 GAGAATTCT 480

RESULT 6
US-10-424-599-38768
; Sequence 38768, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 38768
;   LENGTH: 917
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_135007C.1
```

US-10-424-599-38768

Alignment Scores:
Pred. No.: 1.61e-23 Length: 917
Score: 261.50 Matches: 48
Percent Similarity: 68.13% Conservative: 14
Best Local Similarity: 52.75% Mismatches: 24
Query Match: 49.34% Indels: 5
DB: 16 Gaps: 1

US-10-629-953-4 (1-101) x US-10-424-599-38768 (1-917)

```
Qy         11 GlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSer 30
              |||          |||            :::         ||||||   |||
Db        305 GGCCACGCGCCGCCCATGAAGTCCCTCAAGAAGGATTCCGACGACGATCCC--------- 355

Qy         31 AlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThr 50
                 |||   :::   :::   :::||||||||||||||||||||||||||||||
Db        356 ------GACTCTGTTACCATTTGCGAAATTTCGAAGAACAGGAGGGTTGCCGTGAGGAAC 409

Qy         51 TrpAsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeu 70
              |||   |||   ::::::||||||||||||||||||||||   ||||||||||||   |||
Db        410 TGGAAAGGCAGCATTATGGTTGACATTCGCGAGTTTTACGTCAAAGATGGCAAGCAATTG 469

Qy         71 ProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIle 90
              ||||||||||||||||   |||   ||||||||||||   :::||||||:::::::::
Db        470 CCTGGCAGGAAAGGTATCTCGTTGACCATGGATCAGTGGAATGTGCTTCGTAATCATGTT 529

Qy         91 LysAlaIleAspGluAlaIleLysGluAsnAla 101
              :::  ||||||:::|||:::   ||||||:::
Db        530 GAAGAAATTGACAAGGCAGTTAATGAGAATTCT 562
```

RESULT 7
US-09-770-961-520/c
; Sequence 520, Application US/09770961
; Publication No. US20030115639A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2026 (PARA-015PRV)
; CURRENT APPLICATION NUMBER: US/09/770,961
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,466
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 520
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-961-520

Alignment Scores:
Pred. No.: 6.89e-14 Length: 496
Score: 183.00 Matches: 37
Percent Similarity: 56.25% Conservative: 17
Best Local Similarity: 38.54% Mismatches: 22
Query Match: 34.53% Indels: 20
DB: 10 Gaps: 2

US-10-629-953-4 (1-101) x US-09-770-961-520 (1-496)

```
Qy          4 LysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAlaGly 23
              ||||||||||||   |||
Db        459 AAAGGAAACAAAGAGTTT---------------------------------------- 442

Qy         24 AspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsn 43
              ||||||                  |||   |||:::::::::   :::::|||
Db        441 GATGAT-----------------GACGGCGATCTTATCATTTGCAGGTTATCGGATAAG 400

Qy         44 ArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyr 63
              |||||||||   ::::::   :::   ||||||   :::|||   |||||||||:::|||
Db        399 AGAAGAGTGACGATTCAGGAATTTAAAGGGAAGAGTTTGGTTTCTATCAGAGAGTATTAC 340

Qy         64 GluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrp 83
              :::||||||||||||   ||||||      |||||||||   |||      :::||||||
Db        339 AAGAAAGATGGCAAAGAACTTCCTACTTCTAAAGGAATAAGCTTAACAGATGAACAATGG 280

Qy         84 LysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
                     :::   |||:::   ||||||:::   |||:::|||:::
Db        279 TCAACCTTCAAGAAAAACATGCCAGCCATCGAAAATGCTGTCAAGAAA 232
```

RESULT 8
US-10-425-114-20666
; Sequence 20666, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20666
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-699-H3_FLI
US-10-425-114-20666

Alignment Scores:
Pred. No.: 1.38e-13 Length: 794
Score: 182.50 Matches: 36
Percent Similarity: 63.64% Conservative: 20
Best Local Similarity: 40.91% Mismatches: 23
Query Match: 34.43% Indels: 9
DB: 16 Gaps: 2

US-10-629-953-4 (1-101) x US-10-425-114-20666 (1-794)

```
Qy         12 GlyGlyGluProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAla 31
              ||||||      |||:::|||   :::          |||:::|||
Db        340 GGAGGA------GCTTCCAAGGGCAAGGAGTACGATGATGAAGGC-------------- 378

Qy         32 AspAspAspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrp 51
                    ||:::::::::   ::::::|||       |||||||||   ::::::   :::
Db        379 ------GATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTGACGATTCAGGATTTC 432

Qy         52 AsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuPro 71
                 ||||||   :::|||   |||||||||:::|||:::||||||||||||   ||||||
```

```
Db        433 AGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGATGGCAAGGAACTTCCT 492

Qy         72 GlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLys 91
                 |||||||||   |||      :::||||||         :::   |||:::
Db        493 ACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAGCCTTTAAGAAAAATGTGCCT 552

Qy         92 AlaIleAspGluAlaIleLysGlu 99
              ||||||::::::|||||||||:::
Db        553 GCCATAGAAAAAGCCATTAAGAAA 576

RESULT 9
US-10-424-599-4356
; Sequence 4356, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 4356
;   LENGTH: 911
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_10393C.1
US-10-424-599-4356

Alignment Scores:
Pred. No.:              1.61e-13       Length:         911
Score:                  182.50         Matches:        36
Percent Similarity:     63.64%         Conservative:   20
Best Local Similarity:  40.91%         Mismatches:     23
Query Match:            34.43%         Indels:         9
DB:                     16             Gaps:           2

US-10-629-953-4 (1-101) x US-10-424-599-4356 (1-911)

Qy         12 GlyGlyGluProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAla 31
              ||||||      |||:::|||   :::          |||:::|||
Db        410 GGAGGA------GCTTCCAAGGGCAAGGAGTACGATGATGAAGGC-------------- 448

Qy         32 AspAspAspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrp 51
                 |||:::::::::   :::::|||      |||||||||   ::::::   :::
Db        449 ------GATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGGGTGACGATTCAGGATTTC 502

Qy         52 AsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuPro 71
                 ||||||   :::|||   |||||||||:::|||:::||||||||||||   ||||||
Db        503 AGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAGGATGGCAAGGAACTTCCT 562

Qy         72 GlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLys 91
                 |||||||||   |||      :::||||||         :::   |||:::
Db        563 ACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAGCCTTTAAGAAAAATGTGCCT 622

Qy         92 AlaIleAspGluAlaIleLysGlu 99
              ||||||::::::|||||||||:::
Db        623 GCCATAGAAAAAGCCATTAAGAAA 646

RESULT 10
US-10-437-963-59601/c
; Sequence 59601, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 59601
;   LENGTH: 1030
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_61208C.1
US-10-437-963-59601

Alignment Scores:
Pred. No.:              2.15e-13       Length:         1030
Score:                  182.00         Matches:        41
Percent Similarity:     62.00%         Conservative:   21
Best Local Similarity:  41.00%         Mismatches:     27
Query Match:            34.34%         Indels:         11
DB:                     17             Gaps:           3

US-10-629-953-4 (1-101) x US-10-437-963-59601 (1-1030)

Qy          3 ArgLysGlyAsnLysArgPheGlyGlyGlyGly-GluPro------AlaAlaLysArgAr 20
              |||  |||      |||  ||||||||||| |||                  :::||
Db        770 CGAGGAGGAAGAGGAAGAGGAGGAGGAGGAGGAGGAGCGGGGAGGTGGAACGAAGAG 711

Qy         20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIl 40
              |     |||:::|||                        |||:::::::::   :::::
Db        710 GGAGTACGATGACGAGGGA--------------------GACCTCATCCTCTGCCGCCT 672

Qy         40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
              :|||      |||||||||   ::::::   :::   ||||||   ::::::   |||||
Db        671 TTCGGCCAGGAGGAGGGTGACTCTTCAGGAGTTCAAGGGCAAGACGCTTCTGTCCATCCG 612

Qy         60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMe 80
              ||||:::|||  ||||||||||||  ||||||      |||||||||   |||   ::
Db        611 AGAGTACTACTTCAAGGACGGCAAGGAGCTGCCCGCC---AAAGGGATAAGTTTGACAGT 555

Qy         80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              ::::||||||:::      ||||||::::::  ||||||::::::|||||||||:::
Db        554 TGAGCAGTGGGAAGCATTTCGTGATTCTGTGCCAGCAATAGAGGATGCCATTAAAAAG 497

RESULT 11
US-10-424-599-9157
; Sequence 9157, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9157
;   LENGTH: 1479
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_108279C.1
US-10-424-599-9157

Alignment Scores:
Pred. No.:              7.85e-13       Length:         1479
Score:                  179.00         Matches:        38
```

```
Percent Similarity:     63.53%          Conservative:  16
Best Local Similarity:  44.71%          Mismatches:    28
Query Match:            33.77%          Indels:        3
DB:                     16              Gaps:          1

US-10-629-953-4 (1-101) x US-10-424-599-9157 (1-1479)

Qy         18 LysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAsp-AspAsp------Il 35
              :::||||||          |||   |||   :::::::::   ::: ||||||
Db        661 CGAAGAAGAGTCCCGAGAGACAGAGGTGCTGCCCAAACTGAAGAGGGATGATCCCGAACG 720

Qy         35 eValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysVa 55
              |||:::      :::|||      |||   ||||||||||:::   :::   |||
Db        721 CGTTATTTGCCACCTGTCCAACAGGAGGAACGTGGCGGTGAAAGATTTCAAAGGGACAAC 780

Qy         55 lValValAspIleArgGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGl 75
              :::|||   ||||||||||||||   ||||||||||||   |||||||||   |||||
Db        781 CCTGGTCTCAATTAGGGAGTTCTATATGAAAGATGGAAAACCACTTCCTGGTTCGAAAGG 840

Qy         75 yIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGl 95
              ||||   |||      :::||||||         :::   ::::::   ||||||:::||
Db        841 GATAAGTTTATCTTCTGAACAATGGTCGACCTTCAAGAAGAGTGTTCCTGCCATAGAGGA 900

Qy         95 uAlaIleLysGlu 99
              ||||||||||:::
Db        901 AGCTATCAAAAAG 913

RESULT 12
US-10-424-599-127795/c
; Sequence 127795, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 127795
;   LENGTH: 497
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_86402C.1
US-10-424-599-127795

Alignment Scores:
Pred. No.:              3.97e-13        Length:        497
Score:                  177.00          Matches:       31
Percent Similarity:     66.22%          Conservative:  18
Best Local Similarity:  41.89%          Mismatches:    25
Query Match:            33.40%          Indels:        0
DB:                     16              Gaps:          0

US-10-629-953-4 (1-101) x US-10-424-599-127795 (1-497)

Qy         26 GlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLysAsnArgArg 45
              |||                 :::   |||:::::::::   ::::::|||      ||||||
Db        492 GGCAAGGAGTACGATGATGAAGGCGATCTCATCATCTGCAGGCTTTCAGATAAGAGAAGG 433

Qy         46 ValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyrGluLys 65
              |||   ::::::   :::   ||||||   :::|||   |||||||||:::|||:::|||
Db        432 GTGACGATTCAGGATTTCAGAGGGAAAACATTGGTCTCCATTCGGGAGTATTATAAAAAG 373

Qy         66 AspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIle 85
              |||||||||   ||||||      |||||||||   |||      :::||||||
Db        372 GATGGCAAGGAGCTTCCTACTTCCAAAGGAATAAGTTTGACAGAAGAGCAGTGGTCAACC 313

Qy         86 LeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
```

```
               :::   |||:::   ||||||::::::|||||||||:::
Db        312 TTTAAGAAAAATGTGCCTGCCATAGAAAAAGCCATTAAGAAA 271

RESULT 13
US-10-087-192-1784
; Sequence 1784, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1784
;   LENGTH: 2776
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-087-192-1784

Alignment Scores:
Pred. No.:              1.25e-11        Length:        2776
Score:                  172.00          Matches:       36
Percent Similarity:     64.37%          Conservative:  20
Best Local Similarity:  41.38%          Mismatches:    29
Query Match:            32.45%          Indels:        2
DB:                     13              Gaps:          2

US-10-629-953-4 (1-101) x US-10-087-192-1784 (1-2776)

Qy         14 GluProAlaAlaLysArgArgAlaAlaGlyAspAspGlyProSerGluSerAlaAspAsp 33
              :::|||          ||||||                  |||   |||:::   |||
Db        163 AAGCCTGGTGAGACTTCTAGAGCACTGGCATCCTCCAAGCAGAGCAGCAGCAGCAGAGAT 222

Qy         34 AspIleValValAlaGlnIleSerLysAsnArgArgValAlaValArgThrTrpAsnGly 53
              |||   :::      ||||||   |||   |||   |||:::||||||   :::   |||
Db        223 GACAACATGTTC---CAGATTGGAAAGATGAGATATGTCAGTGTTCGGGACTTCAAAGGA 279

Qy         54 LysValValValAspIleArgGluPheTyr---GluLysAspGlyLysThrLeuProGly 72
              |||:::::::::||||||||||||||::::   :::   :::|||:::      ||||||
Db        280 AAAATTCTAATTGATATTAGAGAATATTGGATGGATTCAGAAGGTGAAATGAAACCAGGA 339

Qy         73 ArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuArgAspAsnIleLysAla 92
              ||||||||||||   |||   |||:::||||||      |||::::::   |||
Db        340 AGAAAAGGTATTTCTTTAAACATGGAACAATGGAGCCAGCTGAAGGAACAGATCTCTGAT 399

Qy         93 IleAspGluAlaIleLysGlu 99
              ||||||:::|||:::::::::
Db        400 ATAGATGACGCAGTAAGAAAG 420

RESULT 14
US-10-425-114-24336
; Sequence 24336, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
```

```
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24336
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3608-031-G8_FLI
US-10-425-114-24336

Alignment Scores:
Pred. No.:              1.76e-11        Length:         933
Score:                  166.50          Matches:        37
Percent Similarity:     53.10%          Conservative:   23
Best Local Similarity:  32.74%          Mismatches:     35
Query Match:            31.42%          Indels:         18
DB:                     16              Gaps:           2

US-10-629-953-4 (1-101) x US-10-425-114-24336 (1-933)

Qy          5 GlyAsnLysArgPheGlyGlyGly-----------GlyGluProAlaAlaLysArgArg 20
              |||   :::      |||||||||           ||||||                 |||
Db        245 GGAGGCGAAGGCGGAGGAGGAGGGCGTCACTGGCAGGGAGAGCAAGGGCAAGGAACACGA 304

Qy         21 AlaAlaGly-AspAspGlyProSerGluSerAlaAspAsp------------------- 33
                 ||| ::::::         |||       |||:::
Db        305 AGAGGAGGACGAAGAGGATGACGATGAGGAGGAGGATGAAGGTAAGGGTGGCGGGAAGAG 364

Qy         34 ------------------AspIleValValAlaGlnIleSerLysAsnArgArgValAl 47
                              |||:::::::::   ::::::|||      |||||||||
Db        365 AGAGTACGATGACCAAGGTGACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGAC 424

Qy         47 aValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGl 67
               :::      :::   |||:::   :::|||   ||||||||||||||   ||||||||
Db        425 TTTATCGGAGTTTAAGGGCAGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGG 484

Qy         67 yLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuAr 87
              ||||   :::|||      |||||||||   :::   ::::::||||||:::
Db        485 CAAGGAGATGCCCTCCGCCAAAGGTATCAGTATGACTTTGGAGCAGTGGGAAGCATTTTG 544

Qy         87 gAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
               :::   :::   ||||||::::::|||||||||:::
Db        545 CAATGCTGTACCTGCAATAGAGGATGCCATCAAAAAG 581

RESULT 15
US-10-425-114-33158
; Sequence 33158, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33158
;   LENGTH: 935
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLMO17069H06_FLI
US-10-425-114-33158

Alignment Scores:
Pred. No.:              1.77e-11        Length:         935
Score:                  166.50          Matches:        37
Percent Similarity:     53.10%          Conservative:   23
Best Local Similarity:  32.74%          Mismatches:     35
Query Match:            31.42%          Indels:         18
DB:                     16              Gaps:           2

US-10-629-953-4 (1-101) x US-10-425-114-33158 (1-935)

Qy          5 GlyAsnLysArgPheGlyGlyGly-----------GlyGluProAlaAlaLysArgArg 20
              |||   :::      |||||||||           ||||||                 |||
Db        263 GGAGGCGAAGGCGGAGGAGGAGGGCGTCACTGGCAGGGAGAGCAAGGGCAAGGAACACGA 322

Qy         21 AlaAlaGly-AspAspGlyProSerGluSerAlaAspAsp------------------- 33
                 ||| ::::::         |||       |||:::
Db        323 AGAGGAGGACGAAGAGGATGACGATGAGGAGGAGGATGAAGGTAAGGGTGGCGGGAAGAG 382

Qy         34 ------------------AspIleValValAlaGlnIleSerLysAsnArgArgValAl 47
                              |||:::::::::   ::::::|||      |||||||||
Db        383 AGAGTACGATGACCAAGGTGACCTTATCCTGTGCCGCCTTTCGAGCAAGAGGAGGGTGAC 442

Qy         47 aValArgThrTrpAsnGlyLysValValValAspIleArgGluPheTyrGluLysAspGl 67
               :::      :::   |||:::   :::|||   ||||||||||||||   ||||||||
Db        443 TTTATCGGAGTTTAAGGGCAGGTCGTTGGTGTCCATCCGCGAGTTCTACGTGAAGGACGG 502

Qy         67 yLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGlnTrpLysIleLeuAr 87
              ||||   :::|||      |||||||||   :::   ::::::||||||:::
Db        503 CAAGGAGATGCCCTCCGCCAAAGGTATCAGTATGACTTTGGAGCAGTGGGAAGCATTTTG 562

Qy         87 gAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
               :::   :::   ||||||::::::|||||||||:::
Db        563 CAATGCTGTACCTGCAATAGAGGATGCCATCAAAAAG 599


Search completed: January 12, 2005, 04:32:28
Job time : 582 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2005, 01:19:56 ; Search time 2548 Seconds
(without alignments)
1444.432 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRFGGGGEPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10629953/runat_10012005_172536_9425/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10629953_@CGN_1_1_3437_@runat_10012005_172536_9425 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	530	100.0	697	6	CB684473	CB684473 OSJNEf14A
c	2	530	100.0	784	6	CB684474	CB684474 OSJNEf14A
	3	517	97.5	693	1	AU082888	AU082888 AU082888
	4	484	91.3	530	7	D15350	D15350 RICC0499A R
	5	425	80.2	523	6	CA116904	CA116904 SCAGLR102
	6	425	80.2	631	6	CA151445	CA151445 SCJFRZ201
	7	425	80.2	635	6	CA230818	CA230818 SCJFFL3C0
	8	425	80.2	637	6	CA259070	CA259070 SCEQRT301
	9	425	80.2	649	6	CA189554	CA189554 SCCCLR1C0
	10	416	78.5	440	7	CF760736	CF760736 DSAF1_65_
	11	415.5	78.4	564	6	CD863080	CD863080 AZO1.105K
	12	415	78.3	855	6	CA129813	CA129813 SCUTLR201
	13	412.5	77.8	532	5	BQ838824	BQ838824 WHE3593_B
	14	409.5	77.3	579	5	BU998906	BU998906 HI12I13r
c	15	409.5	77.3	618	6	CB860066	CB860066 HI12I13w
c	16	409.5	77.3	618	6	CB860077	CB860077 HI12J02w
	17	409.5	77.3	619	5	BU998918	BU998918 HI12J02r
c	18	409.5	77.3	1072	7	CK212042	CK212042 FGAS02390
	19	400.5	75.6	539	1	AU092741	AU092741 AU092741
	20	398	75.1	568	6	CB250145	CB250145 3529_1_18
c	21	398	75.1	596	6	CD568912	CD568912 3529_1_99
	22	396	74.7	470	6	CD913769	CD913769 G550.119C
c	23	395	74.5	603	1	AI691445	AI691445 606018A12
	24	392	74.0	606	6	CA189705	CA189705 SCCCLR1C0
	25	391	73.8	403	2	BE424113	BE424113 WHE0076_G
c	26	391	73.8	618	1	AI734589	AI734589 606033E08
c	27	390	73.6	566	2	BE186543	BE186543 946008A01
	28	389.5	73.5	460	6	CA710341	CA710341 wdk2c.pk0
c	29	381	71.9	623	7	CF635823	CF635823 zmrww00_0
	30	375	70.8	516	6	CF019517	CF019517 QBN1d08.x
	31	373.5	70.5	586	6	CD568913	CD568913 3529_1_99
c	32	363	68.5	588	7	CF625939	CF625939 zmrws05_0
c	33	363	68.5	588	7	CF627159	CF627159 zmrws05_0
	34	363	68.5	652	3	AY112536	AY112536 Zea mays
	35	362.5	68.4	443	6	C28075	C28075 C28075 Rice
	36	347.5	65.6	396	6	C27852	C27852 C27852 Rice
	37	338.5	63.9	555	4	BI795760	BI795760 H030F04 E
	38	329	62.1	288	1	AU068341	AU068341 AU068341
	39	329	62.1	542	5	BU498860	BU498860 946170D05
c	40	299	56.4	485	1	AI943769	AI943769 618045D08
	41	272.5	51.4	586	6	CB079577	CB079577 hp73g03.b
	42	272	51.3	551	6	CB922664	CB922664 VVD086E01
c	43	272	51.3	570	5	BQ796180	BQ796180 EST 5118
	44	272	51.3	601	6	CB339739	CB339739 CA23EI03I
	45	272	51.3	620	6	CB921948	CB921948 VVD077E10

ALIGNMENTS

RESULT 1
CB684473
LOCUS CB684473 697 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEf14A12.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA clone OSJNEf14A12 5', mRNA sequence.
ACCESSION CB684473
VERSION CB684473.1 GI:29688198
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 697)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 14 row: A column: 12
Seq primer: gta aaa cga cgg cca gtg.
FEATURES Location/Qualifiers

```
     source          1. .697
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf14A12"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

Alignment Scores:
Pred. No.:              3.1e-52         Length:         697
Score:                  530.00          Matches:        101
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-10-629-953-4 (1-101) x CB684473 (1-697)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 114

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 174

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 234

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATG 294

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 354

Qy        101 Ala 101
              |||
Db        355 GCG 357

RESULT 2
CB684474/c
LOCUS       CB684474                784 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEf14A12.r OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf14A12 3', mRNA sequence.
ACCESSION   CB684474
VERSION     CB684474.1  GI:29688199
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 14  row: A  column: 12
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf14A12"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

Alignment Scores:
Pred. No.:              3.57e-52        Length:         784
Score:                  530.00          Matches:        101
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-10-629-953-4 (1-101) x CB684474 (1-784)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 ATGTGGCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 581

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 521

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 461

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GAGTTTTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATG 401

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 341

Qy        101 Ala 101
              |||
Db        340 GCG 338

RESULT 3
AU082888
LOCUS       AU082888                693 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AU082888 Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C53889, mRNA sequence.
ACCESSION   AU082888
VERSION     AU082888.1  GI:7212585
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from callus (2000)
  JOURNAL   Unpublished (2000)
```

```
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C53889"
                     /clone_lib="Rice callus"
                     /note="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
                     NotI; cDNA prepared from rice callus mRNAs by using
                     oligo(dT) as a primer and ligating to the SalI-NotI site
                     of pBluescript II SK+ phagemid.  "
ORIGIN

Alignment Scores:
Pred. No.:                1.05e-50      Length:          693
Score:                    517.00        Matches:         100
Percent Similarity:       99.01%        Conservative:    0
Best Local Similarity:    99.01%        Mismatches:      1
Query Match:              97.55%        Indels:          0
DB:                       1             Gaps:            0

US-10-629-953-4 (1-101) x AU082888 (1-693)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ATGTGNCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 121

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 181

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 241

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATG 301

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 361

Qy        101 Ala 101
              |||
Db        362 GCG 364

RESULT 4
D15350
LOCUS       D15350                   530 bp    mRNA    linear   EST 24-FEB-2003
DEFINITION  RICC0499A Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C0499, mRNA sequence.
ACCESSION   D15350 C98059
VERSION     D15350.2  GI:15072274
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Sasaki,T., Song,J., Koga-Ban,Y., Matsui,E., Fang,F., Higo,H.,
            Nagasaki,H., Hori,M., Miya,M. and Murayama-Kayano,E.
  TITLE     Toward cataloguing all rice genes: large-scale sequencing of
            randomly chosen rice cDNAs from a callus cDNA library
  JOURNAL   Plant J. 6 (4), 615-624 (1994)
  MEDLINE   95078950
   PUBMED   7987417
COMMENT     On May 17, 1993 this sequence version replaced gi:286543
            gi:3760805.
            Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'
            Seq primer: oligo(dT).
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C0499"
                     /clone_lib="Rice callus"
                     /note="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
                     NotI; cDNA prepared from rice callus mRNAs by using
                     oligo(dT) as a primer and ligating to the SalI-NotI site
                     of pBluescript II SK+ phagemid.  "
ORIGIN

Alignment Scores:
Pred. No.:                5.95e-47      Length:          530
Score:                    484.00        Matches:         96
Percent Similarity:       96.00%        Conservative:    0
Best Local Similarity:    96.00%        Mismatches:      3
Query Match:              91.32%        Indels:          1
DB:                       7             Gaps:            0

US-10-629-953-4 (1-101) x D15350 (1-530)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGly-GlyGlyGlyGluProAlaAlaLysArgAr 20
              |||  |||||||||||||||||||||||| |||||||||   |||   ||||||||||||
Db         66 ATGNGGCGGAAGGGGAACAAGCGGTTCGGACGGCGGCGGCGANCCGACGGCCAAGCGCCG 125

Qy         20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIl 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTNGCCCAGAT 185

Qy         40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCG 245

Qy         60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGAT 305

Qy         80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAG 363

RESULT 5
CA116904
LOCUS       CA116904                 523 bp    mRNA    linear   EST 23-SEP-2003
DEFINITION  SCAGLR1021A08.g LR1 Saccharum officinarum cDNA clone SCAGLR1021A08
            5', mRNA sequence.
ACCESSION   CA116904
VERSION     CA116904.1  GI:34970211
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
```

```
REFERENCE   1  (bases 1 to 523)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 021  row: A  column: 08
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCAGLR1021A08"
                     /lab_host="DH10B"
                     /clone_lib="LR1"
                     /note="Organ: Leaf roll from field grown adult plants
                     (large insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Leaf roll from field grown  adult plants (large
                     insert library)]. cDNA was prepared from polyA+ mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:              5.33e-40        Length:         523
Score:                  425.00          Matches:        79
Percent Similarity:     89.00%          Conservative:   10
Best Local Similarity:  79.00%          Mismatches:     11
Query Match:            80.19%          Indels:         0
DB:                     6               Gaps:           0

US-10-629-953-4 (1-101) x CA116904 (1-523)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||   ||||||   ||||||||||||||||||||||||||||||||||||||:::
Db         12 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCAG 71

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||    |||:::|||       |||||||||:::|||
Db         72 GCCGCCGGGGATGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 131

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||::::::|||:::||||||:::|||   |||:::|||   |||||:::|||
Db        132 TCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTACGTCGACCTCCGC 191

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   ||||||||||||||||||||   ||||||||||||   |||   :::
Db        192 GAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTCCAGTTA 251

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGATGAGGCCATCAAGGAGAAC 311

RESULT 6
CA151445
LOCUS       CA151445                631 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCJFRZ2010C02.g RZ2 Saccharum officinarum cDNA clone SCJFRZ2010C02
            5', mRNA sequence.
ACCESSION   CA151445
VERSION     CA151445.1  GI:35055505
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 010  row: C  column: 02
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCJFRZ2010C02"
                     /lab_host="DH10B"
                     /clone_lib="RZ2"
                     /note="Organ: Shoot-root transition zone from young plants
                     (small insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Shoot-root transition zone  from young plants (small
                     insert library)]. cDNA was prepared from polyA+  mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:              6.67e-40        Length:         631
Score:                  425.00          Matches:        79
Percent Similarity:     89.00%          Conservative:   10
Best Local Similarity:  79.00%          Mismatches:     11
Query Match:            80.19%          Indels:         0
DB:                     6               Gaps:           0

US-10-629-953-4 (1-101) x CA151445 (1-631)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||   ||||||   ||||||||||||||||||||||||||||||||||||||:::
Db          9 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCAG 68

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||    |||:::|||       |||||||||:::|||
Db         69 GCCGCCGGGGATGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 128

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||::::::|||:::||||||:::|||   |||:::|||   |||||:::|||
Db        129 TCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAGGGCAGGGTCTACGTCGACCTCCGC 188

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   ||||||||||||||||||||   ||||||||||||   |||   :::
Db        189 GAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTCCAGTTA 248

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        249 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCCATAGATGAGGCCATCAAGGAGAAC 308

RESULT 7
CA230818
LOCUS       CA230818                635 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCJFFL3C07F05.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCJFFL3C07F05 5', mRNA sequence.
ACCESSION   CA230818
VERSION     CA230818.1  GI:35294733
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: C07  row: F  column: 05
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCJFFL3C07F05"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing  inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:              6.73e-40        Length:         635
Score:                  425.00          Matches:        79
Percent Similarity:     89.00%          Conservative:   10
Best Local Similarity:  79.00%          Mismatches:     11
Query Match:            80.19%          Indels:         0
DB:                     6               Gaps:           0

US-10-629-953-4 (1-101) x CA230818 (1-635)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||  ||||||  ||||||||||||||||||||||||||||||||||||||||||:::
Db         19 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCAG 78

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||  |||:::|||      |||||||||:::|||
Db         79 GCCGCCGGGGATGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 138

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||::::::|||:::||||||:::|||  |||:::|||  ||||||:::|||

Db        139 TCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTACGTCGACCTCCGC 198

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||  ||||||||||||||||||||  ||||||||||||  |||   :::
Db        199 GAGTTCTACGTCAAGGACGGCAAGACTCTCCCCCACCCGCAAAGGTATATCACTCCAGTTA 258

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGACGAGGCCATCAAGGAGAAC 318

RESULT 8
CA259070
LOCUS       CA259070                637 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  SCEQRT3016B12.g RT3 Saccharum officinarum cDNA clone SCEQRT3016B12
            5', mRNA sequence.
ACCESSION   CA259070
VERSION     CA259070.1  GI:35947628
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 016  row: B  column: 12
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCEQRT3016B12"
                     /lab_host="DH10B"
                     /clone_lib="RT3"
                     /note="Organ: Root apex from adult plants; Vector:
                     pSport1; Site_1: SalI; Site_2: NotI; An unidirectional
                     cDNA library generated from [Root apex from adult
                     plants]. cDNA was prepared from polyA+ mRNA using
                     SuperScript Plasmid  System Kit (Invitrogen). The
                     double-strand cDNAs were fractionated in a  sepharose
                     CL-2B 40cm-columns and fragments sizing between 0.8 and
                     1.5 Kb  were directionally cloned into the vector. Details
                     of each source of RNA  and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:              6.75e-40        Length:         637
Score:                  425.00          Matches:        79
Percent Similarity:     89.00%          Conservative:   10
Best Local Similarity:  79.00%          Mismatches:     11
Query Match:            80.19%          Indels:         0
DB:                     6               Gaps:           0

US-10-629-953-4 (1-101) x CA259070 (1-637)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||  ||||||  ||||||||||||||||||||||||||||||||||||||||||:::
Db         18 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGTGGCGGCGGCGAGCCGGCGGCCAAGCGCCAG 77
```

```
Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||    |||:::|||       |||||||||:::|||
Db         78 GCCGCCGGGGACGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 137

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||:::::::|||:::|||||:::|||   |||:::|||   |||||:::|||
Db        138 TCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTACGTCGACCTCCGC 197

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   |||||||||||||||||||   |||||||||||   |||   :::
Db        198 GAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTACAGTTA 257

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGACGAGGCCATCAAGGAGAAC 317

RESULT 9
CA189554
LOCUS       CA189554                649 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCCCLR1C02A09.g LR1 Saccharum officinarum cDNA clone SCCCLR1C02A09
            5', mRNA sequence.
ACCESSION   CA189554
VERSION     CA189554.1  GI:35132819
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: C02  row: A  column: 09
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCCCLR1C02A09"
                     /lab_host="DH10B"
                     /clone_lib="LR1"
                     /note="Organ: Leaf roll from field grown adult plants
                     (large insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Leaf roll from field grown  adult plants (large
                     insert library)]. cDNA was prepared from polyA+ mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:              6.9e-40        Length:        649
Score:                  425.00         Matches:       79
Percent Similarity:     89.00%         Conservative:  10
Best Local Similarity:  79.00%         Mismatches:    11

Query Match:            80.19%         Indels:        0
DB:                     6              Gaps:          0

US-10-629-953-4 (1-101) x CA189554 (1-649)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||   ||||||   ||||||||||||||||||||||||||||||||||||:::
Db         12 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCAG 71

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||    |||:::|||       |||||||||:::|||
Db         72 GCCGCCGGGGATGACGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 131

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||:::::|||:::|||||:::|||   |||:::|||   |||||:::|||
Db        132 TCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTACGTCGACCTCCGC 191

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   |||||||||||||||||||   |||||||||||   |||   :::
Db        192 GAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTCCAGTTA 251

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGATGAGGCCATCAAGGAGAAC 311

RESULT 10
CF760736
LOCUS       CF760736                440 bp    mRNA    linear   EST 17-OCT-2003
DEFINITION  DSAF1_65_F02.b1_A011 Drought-stressed after flowering Sorghum
            bicolor cDNA clone DSAF1_65_F02_A011 5', mRNA sequence.
ACCESSION   CF760736
VERSION     CF760736.1  GI:37709961
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Cordonnier-Pratt,M.-M., Zhang,D., McCartor,K., Nguyen,H.T. and
            Pratt,L.H.
  TITLE     An EST Database from Sorghum: Subtracted post-flowering drought
            stressed leaf tissues
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed at Texas Tech University by Deshui Zhang and
            Jianhang Jia in the laboratory of Dr. Henry Nguyen. Sequencing was
            done in the Laboratory for Genomics and Bioinformatics, University
            of Georgia. Sequence ends have been trimmed to exclude vector and
            regions below Phred quality 16. Three-prime sequences are presented
            as their reverse complement and have been trimmed to exclude polyA.
            Seq primer: JENREV (CAGGAAACAGCTATGACC).
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="B35"
                     /db_xref="taxon:4558"
                     /clone="DSAF1_65_F02_A011"
                     /dev_stage="Post-flowering"
                     /lab_host="ElectroMax DH10B (BRL)"
                     /clone_lib="Drought-stressed after flowering"
                     /note="Organ: Leaf; Vector: pBlueScriptSK-; Site_1: XhoI;
                     Site_2: EcoRI; The library was prepared from polyA+ RNA
                     from leaves harvested from post-flowering,
                     drought-stressed Sorghum bicolor, cv. B35. Double-stranded
                     cDNA was cloned unidirectionally using the UniZap system
```

from Stratagene. After amplification, the library was subtracted by re-association hybridization. Inserts can be excised with XhoI and EcoRI."
ORIGIN

Alignment Scores:

Pred. No.:	5e-39	Length:	440
Score:	416.00	Matches:	77
Percent Similarity:	88.00%	Conservative:	11
Best Local Similarity:	77.00%	Mismatches:	12
Query Match:	78.49%	Indels:	0
DB:	7	Gaps:	0

US-10-629-953-4 (1-101) x CF760736 (1-440)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||   ||||||   ||||||||||||||||||||||||||||||||||||||:::
Db         10 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGGCAG 69

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||:::|||||||||||||   |||:::|||       |||||||||:::|||
Db         70 GCCGCCGGGGACGAAGGCCCCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 129

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||:::::|||:::||||||   |||   |||:::|||   ||||||:::|||
Db        130 TCGAAGAACAAGAAGGTGTCTGTTAGGAACTGGAAAGGCAGGGTCTTCGTTGACCTCCGC 189

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   ||||||||||:::||||||   ||||||||||   |||   :::
Db        190 GAGTTCTACGTCAAGGACGGCAAGTCTCTCCCCACCCGCAAAGGTATATCACTCCAGTTA 249

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGATGAGGCCATCAAGGAGAAC 309
```

RESULT 11
CD863080

LOCUS	CD863080	564 bp	mRNA	linear	EST 11-JUL-2003

DEFINITION AZO1.105K01F010129 AZO1 Triticum aestivum cDNA clone AZO1105K01, mRNA sequence.
ACCESSION CD863080
VERSION CD863080.1 GI:32546896
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 564)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).
FEATURES Location/Qualifiers
source 1. .564
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZO1105K01"
/tissue_type="leaf"
/clone_lib="AZO1"
ORIGIN

Alignment Scores:

Pred. No.:	7.7e-39	Length:	564
Score:	415.50	Matches:	80
Percent Similarity:	90.82%	Conservative:	9
Best Local Similarity:	81.63%	Mismatches:	8
Query Match:	78.40%	Indels:	1
DB:	6	Gaps:	1

US-10-629-953-4 (1-101) x CD863080 (1-564)

```
Qy          4 LysGlyAsnLysArgPhe---GlyGlyGlyGlyGluProAlaAlaLysArgArgAlaAla 22
              :::|||||||||||||||   |||||||||||:::|||   ||||||||||:::||||||
Db         23 CGGGGGAACAAGCGCTTCGGCGGCGGCGGCGGCCAGCCGCCGGCCAAGCGCCAGGCGGCC 82

Qy         23 GlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIleSerLys 42
              |||   ||||||||||||||||       ||||||   |||||||||||||||||||||||
Db         83 GGGAAGGACGGCCCCTCCGAGGAAACCGACGACGGCATCGTCGTTGCCCAGATATCGAAG 142

Qy         43 AsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArgGluPhe 62
              |||:::|||||||||||||||:::|||||||||||||||||:::|||||||:::||||||||||
Db        143 AATAAGAGGGTGGCTGTGAGGAGCTGGAACGGCAAGGTGATGGTCGACATGCGCGAGTTC 202

Qy         63 TyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMetAspGln 82
              ||||||||||||||||||:::||||||   ||||||||||   |||   ||||||||||
Db        203 TACGAAAAGGACGGCAAGAGCCTCCCGACCCGCAAAGGTATATCGCTCTCAATGGATCAG 262

Qy         83 TrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db        263 TGGAAAATACTGAGGGACAACATCGAAGCTATAGACGAGGCCATCAAGGAGAAC 316
```

RESULT 12
CA129813

LOCUS	CA129813	855 bp	mRNA	linear	EST 24-SEP-2003

DEFINITION SCUTLR2015B02.g LR2 Saccharum officinarum cDNA clone SCUTLR2015B02 5', mRNA sequence.
ACCESSION CA129813
VERSION CA129813.1 GI:35012058
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE 1 (bases 1 to 855)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br
Plate: 015 row: B column: 02
Seq primer: T7 Promoter Primer.
FEATURES Location/Qualifiers
source 1. .855
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUTLR2015B02"
/lab_host="DH10B"
/clone_lib="LR2"
/note="Organ: Leaf roll from field grown adult plants (small insert library); Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Leaf roll from field grown adult plants (small insert library)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose

CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:

Pred. No.:	1.45e-38	Length:	855
Score:	415.00	Matches:	78
Percent Similarity:	88.00%	Conservative:	10
Best Local Similarity:	78.00%	Mismatches:	12
Query Match:	78.30%	Indels:	0
DB:	6	Gaps:	0

US-10-629-953-4 (1-101) x CA129813 (1-855)

```
Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||  ||||||  |||||||||||||||||||||||||||||||||||||||:::
Db          6 ATGTGGGGGAAGGGAAAGAAGCGTTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCAG 65

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||  ||||||  |||:::|||      |||||||||:::|||
Db         66 GCCGCCGGGGACGACGGCCTCTCCGAACCCGCCGAAGACGGTACGGTCGTAGCCGAGATA 125

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||::::::|||:::||||||:::|||   |||:::|||   ||||||:::|||
Db        126 TCGAAGAACAAGAAGGTGTCCGTTAGGAGCTGGAAAGGCAGGGTCTACGTCGACCTCCGC 185

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   ||||||||||||||||||||   ||||||||||||   |||  :::
Db        186 GAGTTCTACGTCAAGGACGGCAAGACTCTCCCCACCCGCAAAGGTATATCACTACAGTTA 245

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GATCAGTGGAAGATATTGAGGGACAATATCAAAGCCATAGACGAGGCCATCAAGGAGAAC 305
```

RESULT 13
BQ838824

LOCUS	BQ838824	532 bp	mRNA	linear	EST 08-AUG-2002

DEFINITION WHE3593_B06_D11ZS Wheat developing grains cDNA library Triticum aestivum cDNA clone WHE3593_B06_D11, mRNA sequence.
ACCESSION BQ838824
VERSION BQ838824.1 GI:22143146
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 532)
AUTHORS Altenbach,S., Anderson,O.D., Chao,S., Chin,A., Close,T.J., Cronin,K., Crossman,C., Fenton,R.D., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
TITLE The structure and function of the expressed portion of the wheat genomes - Developing grains cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES Location/Qualifiers
source 1. .532
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Butte 86"
/db_xref="taxon:4565"
/clone="WHE3593_B06_D11"
/tissue_type="whole grains"
/dev_stage="3-44 days post anthesis seed"
/lab_host="E. coli SOLR"
/clone_lib="Wheat developing grains cDNA library"
/note="Vector: Lambda ZAP II, excised phagemid; Site_1: EcoRI; Plants were grown under six following different environmental regimes in greenhouse, Environment 1) 24oC/17oC day/night, well-watered, with post-anthesis fertilizer, Environment 2) 24oC/17oC day/night, well-watered, without post-anthesis fertilizer, Environment 3) 37oC/17oC day/night, well-watered, with post-anthesis fertilizer, Environment 4) 37oC/17oC day/night, well-watered, without post-anthesis fertilizer, Environment 5) 37oC/17oC day/night plus drought, with post-anthesis fertilizer, Environment 6) 37oC/17oC day/night plus drought, without post-anthesis fertilizer, developing wheat grains from the following were excised and frozen in liquid nitrogen, Environment 1 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44 DPA Environment 2 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44 DPA Environment 3 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 34 DPA Environment 4 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 34 DPA Environment 5 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 30 DPA Environment 6 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 30 DPA and total RNA was prepared by S. Altenbach and K. Cronin at USDA-ARS, Albany, CA. A cDNA library was made using poly (A) RNA, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in the TJ Close lab (Chin, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (others)."

ORIGIN

Alignment Scores:

Pred. No.:	1.62e-38	Length:	532
Score:	412.50	Matches:	80
Percent Similarity:	88.00%	Conservative:	8
Best Local Similarity:	80.00%	Mismatches:	9
Query Match:	77.83%	Indels:	3
DB:	5	Gaps:	1

US-10-629-953-4 (1-101) x BQ838824 (1-532)

```
Qy          4 LysGlyAsnLysArgPheGlyGlyGly---------GlyGluProAlaAlaLysArgArg 20
              :::|||||||||||||||||||||||         |||:::|||   |||||||:::
Db          3 CAGGGGAACAAGCGCTTCGGCGGCGGCGGCGGCCGCGGCCAGCCGCCGGCCAAGCGCCAG 62

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              |||||||||   ||||||||||||||||      ||||||   ||||||||||||||||||
Db         63 GCGGCCGGGAAGGACGGTCCCTCCGAGGAAACCGACGACGGCATCGTCGTTGCGCAGATA 122

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||:::||||||||||||||   |||||||||||||||:::||||||:::|||
Db        123 TCGAAGAACAAGAGGGTGGCCGTGAGGAACTGGAACGGGAAGGTCATGGTCGACATGCGC 182

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||:::||||||   ||||||||||||   |||   |||
Db        183 GAGTTCTACGAAAAGGACGGCAAGAGCCTCCCGACCCGCAAAGGTATATCGCTCTCAATG 242

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        243 GATCAGTGGAAAATACTGAGGGACAACATCGAAGCTATAGACGAGGCCATCAAGGAGAAC 302
```

RESULT 14
BU998906

LOCUS	BU998906	579 bp	mRNA	linear	EST 23-OCT-2002

DEFINITION HI12I13r HI Hordeum vulgare subsp. vulgare cDNA clone HI12I13 5-PRIME, mRNA sequence.
ACCESSION BU998906
VERSION BU998906.1 GI:24275889
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare

```
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Zhang,H., Weschke,W., Michalek,W., Stein,N. and Graner,A.
  TITLE     EST sequencing and analysis in barley (2002)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Stein Nils
            Molecular Markers Group, Department Genbank
            Institute of Plant Genetics and Crop Plant Research (IPK)
            Corrensstr. 3, 06466, Gatersleben, Germany
            Tel: 039482-5522
            Fax: 039482-5595
            Email: stein@ipk-gatersleben.de
            Insert Length: 579   Std Error: 0.00
            Plate: 12  row: I  column: 13
            Seq primer: M13rev.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /sub_species="vulgare"
                     /db_xref="GABI:252770"
                     /db_xref="taxon:112509"
                     /clone="HI12I13"
                     /tissue_type="female inflorescences"
                     /dev_stage="female inflorescences (approx. 3 mm in size)"
                     /lab_host="XL10-Gold"
                     /clone_lib="HI"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of
                     cDNA); Site_2: XhoI (3'-end of cDNA); Due to a cloning
                     artefact caused by the kit, in most cases the EcoRI site
                     is NOT present, as well as the EcoRIadapter used for
                     cloning. To excise the insert, restriction sites upstream
                     EcoRI should be used (e.g. BamHI, SalI,PstI). NOTE: Also
                     due to the cloning system used Blue/white selection for
                     recombinats is not 100% reliable."
ORIGIN

Alignment Scores:
Pred. No.:              4.06e-38        Length:         579
Score:                  409.50          Matches:        79
Percent Similarity:     88.00%          Conservative:   9
Best Local Similarity:  79.00%          Mismatches:     9
Query Match:            77.26%          Indels:         3
DB:                     5               Gaps:           1

US-10-629-953-4 (1-101) x BU998906 (1-579)

Qy          4 LysGlyAsnLysArgPhe---------GlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||         ||||||||||||:::|||   |||||||||:::
Db         14 AAGGGGAACAAGCGCTTCGGCGGCGGCGGCGGTGGCGGCCAGCCGCCGGCCAAGCGCCAG 73

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              |||||||||   ||||||||||||||      ||||||   ||||||||||||||||||
Db         74 GCGGCCGGGAAGGACGGCCCCTCCGAGGAAACCGACGACGGCATCGTCGTCGCGCAGATA 133

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||:::|||||||||||||||:::|||||||||||||||:::||||||:::|||
Db        134 TCGAAGAACAAGAGGGTGGCTGTGCGGAGCTGGAATGGCAAGGTCATGGTCGACATGCGC 193

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   ||||||||||:::|||||   ||||||||||   |||   |||
Db        194 GAGTTCTACGTCAAGGACGGCAAGAGCCTCCCGACCCGCAAAGGTATATCGCTCTCAATG 253

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              |||||||||:::||||||||||||||||||||||||||||||||:::|||||||||
Db        254 GATCAGTGGGAGATACTGAGGGACAACATCAAAGCCATAGACGAGGCCGTCAAGGAGAAC 313

RESULT 15
CB860066/c
```

```
LOCUS       CB860066               618 bp    mRNA    linear   EST 22-APR-2003
DEFINITION  HI12I13w HI Hordeum vulgare subsp. vulgare cDNA clone HI12I13
            3-PRIME, mRNA sequence.
ACCESSION   CB860066
VERSION     CB860066.1  GI:30054625
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Zhang,H., Weschke,W., Michalek,W., Stein,N. and Graner,A.
  TITLE     EST sequencing and analysis in barley (2002)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Stein Nils
            Molecular Markers Group, Department Genbank
            Institute of Plant Genetics and Crop Plant Research (IPK)
            Corrensstr. 3, 06466, Gatersleben, Germany
            Tel: 039482-5522
            Fax: 039482-5595
            Email: stein@ipk-gatersleben.de
            Insert Length: 618   Std Error: 0.00
            Plate: 12  row: I  column: 13
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /sub_species="vulgare"
                     /db_xref="GABI:557649"
                     /db_xref="taxon:112509"
                     /clone="HI12I13"
                     /tissue_type="female inflorescences"
                     /dev_stage="female inflorescences (approx. 3 mm in size)"
                     /lab_host="XL10-Gold"
                     /clone_lib="HI"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of
                     cDNA); Site_2: XhoI (3'-end of cDNA); Due to a cloning
                     artefact caused by the kit, in most cases the EcoRI site
                     is NOT present, as well as the EcoRIadapter used for
                     cloning. To excise the insert, restriction sites upstream
                     EcoRI should be used (e.g. BamHI, SalI,PstI). NOTE: Also
                     due to the cloning system used Blue/white selection for
                     recombinats is not 100% reliable."
ORIGIN

Alignment Scores:
Pred. No.:              4.38e-38        Length:         618
Score:                  409.50          Matches:        79
Percent Similarity:     88.00%          Conservative:   9
Best Local Similarity:  79.00%          Mismatches:     9
Query Match:            77.26%          Indels:         3
DB:                     6               Gaps:           1

US-10-629-953-4 (1-101) x CB860066 (1-618)

Qy          4 LysGlyAsnLysArgPhe---------GlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||||||||||||||||||         ||||||||||||:::|||   |||||||||:::
Db        607 AAGGGGAACAAGCGCTTCGGCGGCGGCGGCGGTGGCGGCCAGCCGCCGGCCAAGCGCCAG 548

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              |||||||||   ||||||||||||||      ||||||   ||||||||||||||||||
Db        547 GCGGCCGGGAAGGACGGCCCCTCCGAGGAAACCGACGACGGCATCGTCGTCGCGCAGATA 488

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              |||||||||:::|||||||||||||||:::|||||||||||||||:::||||||:::|||
Db        487 TCGAAGAACAAGAGGGTGGCTGTGCGGAGCTGGAATGGCAAGGTCATGGTCGACATGCGC 428

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              |||||||||   ||||||||||:::|||||   ||||||||||   |||   |||
Db        427 GAGTTCTACGTCAAGGACGGCAAGAGCCTCCCGACCCGCAAAGGTATATCGCTCTCAATG 368
```

```
Qy          81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
               ||||||||:::||||||||||||||||||||||||||||||||:::|||||||||
Db         367 GATCAGTGGGAGATACTGAGGGACAACATCAAAGCCATAGACGAGGCCGTCAAGGAGAAC 308
```

Search completed: January 12, 2005, 03:27:53
Job time : 2553 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2005, 01:19:56 ; Search time 2548 Seconds
(without alignments)
1444.432 Million cell updates/sec

Title: US-10-629-953-4
Perfect score: 530
Sequence: 1 MWRKGNKRPGGGGBPAAKRR.....QWKILRDNIKAIDEAIKENA 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10629953/runat_10012005_172536_9425/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10629953_@CGN_1_1_3437_@runat_10012005_172536_9425 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	530	100.0	697	6	CB684473	CB684473 OSJNEf14A
c	2	530	100.0	784	6	CB684474	CB684474 OSJNEf14A
	3	517	97.5	693	1	AU082888	AU082888 AU082888
	4	484	91.3	530	7	D15350	D15350 RICC0499A R
	5	425	80.2	523	6	CA116904	CA116904 SCAGLR102
	6	425	80.2	631	6	CA151445	CA151445 SCJFRZ201
	7	425	80.2	635	6	CA230818	CA230818 SCJFFL3C0
	8	425	80.2	637	6	CA259070	CA259070 SCEQRT301
	9	425	80.2	649	6	CA189554	CA189554 SCCCLR1C0
	10	416	78.5	440	7	CF760736	CF760736 DSAF1_65_
	11	415.5	78.4	564	6	CD863080	CD863080 AZO1.105K
	12	415	78.3	855	6	CA129813	CA129813 SCUTLR201
	13	412.5	77.8	532	5	BQ838824	BQ838824 WHE3593_B
	14	409.5	77.3	579	5	BU998906	BU998906 HI12I13r
c	15	409.5	77.3	618	6	CB860066	CB860066 HI12I13w
c	16	409.5	77.3	618	6	CB860077	CB860077 HI12J02w
	17	409.5	77.3	619	5	BU998918	BU998918 HI12J02r
c	18	409.5	77.3	1072	7	CK212042	CK212042 FGAS02390
	19	400.5	75.6	539	1	AU092741	AU092741 AU092741
	20	398	75.1	568	6	CB250145	CB250145 3529_1_18
c	21	398	75.1	596	6	CD568912	CD568912 3529_1_99
	22	396	74.7	470	6	CD913769	CD913769 G550.119C
c	23	395	74.5	603	1	AI691445	AI691445 606018A12
	24	392	74.0	606	6	CA189705	CA189705 SCCCLR1C0
	25	391	73.8	403	2	BE424113	BE424113 WHE0076_G
c	26	391	73.8	618	1	AI734589	AI734589 606033E08
c	27	390	73.6	566	2	BE186543	BE186543 946008A01
	28	389.5	73.5	460	6	CA710341	CA710341 wdk2c.pk0
c	29	381	71.9	623	7	CF635823	CF635823 zmrww00_0
	30	375	70.8	516	6	CF019517	CF019517 QBN1d08.x
	31	373.5	70.5	586	6	CD568913	CD568913 3529_1_99
c	32	363	68.5	588	7	CF625939	CF625939 zmrws05_0
c	33	363	68.5	588	7	CF627159	CF627159 zmrws05_0
	34	363	68.5	652	3	AY112536	AY112536 Zea mays
	35	362.5	68.4	443	6	C28075	C28075 C28075 Rice
	36	347.5	65.6	396	6	C27852	C27852 C27852 Rice
	37	338.5	63.9	555	4	BI795760	BI795760 H030F04 E
	38	329	62.1	288	1	AU068341	AU068341 AU068341
	39	329	62.1	542	5	BU498860	BU498860 946170D05
c	40	299	56.4	485	1	AI943769	AI943769 618045D08
	41	272.5	51.4	586	6	CB079577	CB079577 hp73g03.b
	42	272	51.3	551	6	CB922664	CB922664 VVD086E01
c	43	272	51.3	570	5	BQ796180	BQ796180 EST 5118
	44	272	51.3	601	6	CB339739	CB339739 CA23EI03I
	45	272	51.3	620	6	CB921948	CB921948 VVD077E10

ALIGNMENTS

RESULT 1
CB684473
LOCUS CB684473 697 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEf14A12.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA clone OSJNEf14A12 5', mRNA sequence.
ACCESSION CB684473
VERSION CB684473.1 GI:29688198
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 697)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 14 row: A column: 12
Seq primer: gta aaa cga cgg cca gtg.
FEATURES Location/Qualifiers

THIS PAGE BLANK (USPTO)

```
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES              Location/Qualifiers
     source           1. .693
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="mRNA"
                      /cultivar="Nipponbare"
                      /db_xref="taxon:39947"
                      /clone="C53889"
                      /clone_lib="Rice callus"
                      /note="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
                      NotI; cDNA prepared from rice callus mRNAs by using
                      oligo(dT) as a primer and ligating to the SalI-NotI site
                      of pBluescript II SK+ phagemid.  "
ORIGIN

Alignment Scores:
Pred. No.:                1.05e-50        Length:          693
Score:                    517.00          Matches:         100
Percent Similarity:       99.01%          Conservative:    0
Best Local Similarity:    99.01%          Mismatches:      1
Query Match:              97.55%          Indels:          0
DB:                       1               Gaps:            0

US-10-629-953-4 (1-101) x AU082888 (1-693)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGlyGlyGlyGlyGluProAlaAlaLysArgArg 20
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ATGTGNCGGAAGGGGAACAAGCGGTTCGGCGGCGGCGGCGAGCCGGCGGCCAAGCGCCGT 121

Qy         21 AlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIle 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTCGCCCAGATA 181

Qy         41 SerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCGC 241

Qy         61 GluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGATG 301

Qy         81 AspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGluAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAGAAT 361

Qy        101 Ala 101
              |||
Db        362 GCG 364

RESULT 4
D15350
LOCUS       D15350                  530 bp    mRNA    linear   EST 24-FEB-2003
DEFINITION  RICC0499A Rice callus Oryza sativa (japonica cultivar-group) cDNA
            clone C0499, mRNA sequence.
ACCESSION   D15350 C98059
VERSION     D15350.2  GI:15072274
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Sasaki,T., Song,J., Koga-Ban,Y., Matsui,E., Pang,P., Higo,H.,
            Nagasaki,H., Hori,M., Miya,M. and Murayama-Kayano,E.
  TITLE     Toward cataloguing all rice genes: large-scale sequencing of
            randomly chosen rice cDNAs from a callus cDNA library
  JOURNAL   Plant J. 6 (4), 615-624 (1994)
  MEDLINE   95078950
   PUBMED   7987417
COMMENT     On May 17, 1993 this sequence version replaced gi:286543
            gi:3760805.
            Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'
            Seq primer: oligo(dT).
FEATURES              Location/Qualifiers
     source           1. .530
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="mRNA"
                      /cultivar="Nipponbare"
                      /db_xref="taxon:39947"
                      /clone="C0499"
                      /clone_lib="Rice callus"
                      /note="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
                      NotI; cDNA prepared from rice callus mRNAs by using
                      oligo(dT) as a primer and ligating to the SalI-NotI site
                      of pBluescript II SK+ phagemid.  "
ORIGIN

Alignment Scores:
Pred. No.:                5.95e-47        Length:          530
Score:                    484.00          Matches:         96
Percent Similarity:       96.00%          Conservative:    0
Best Local Similarity:    96.00%          Mismatches:      3
Query Match:              91.32%          Indels:          1
DB:                       2               Gaps:            0

US-10-629-953-4 (1-101) x D15350 (1-530)

Qy          1 MetTrpArgLysGlyAsnLysArgPheGly-GlyGlyGlyGluProAlaAlaLysArgAr 20
              ||| |||||||||||||||||||||||| |||||||||    |||   ||||||||||
Db         66 ATGNGGCGGAAGGGGAACAAGCGGTTCGGACGGCGGCGGCGANCCGACGGCCAAGCGCCG 125

Qy         20 gAlaAlaGlyAspAspGlyProSerGluSerAlaAspAspAspIleValValAlaGlnIl 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TGCCGCCGGGGACGACGGGCCCTCCGAGAGCGCCGACGACGATATCGTCGTNGCCCAGAT 185

Qy         40 eSerLysAsnArgArgValAlaValArgThrTrpAsnGlyLysValValValAspIleAr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATCGAAGAACAGGAGGGTGGCGGTGCGGACCTGGAACGGCAAGGTCGTCGTCGACATCCG 245

Qy         60 gGluPheTyrGluLysAspGlyLysThrLeuProGlyArgLysGlyIleGlnLeuProMe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CGAGTTCTACGAGAAGGACGGCAAGACCCTCCCCGGCCGCAAAGGTATACAGCTCCCGAT 305

Qy         80 tAspGlnTrpLysIleLeuArgAspAsnIleLysAlaIleAspGluAlaIleLysGlu 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GGATCAGTGGAAGATACTGAGGGACAATATCAAAGCTATAGATGAGGCCATCAAGGAG 363

RESULT 5
CA116904
LOCUS       CA116904                523 bp    mRNA    linear   EST 23-SEP-2003
DEFINITION  SCAGLR1021A08.g LR1 Saccharum officinarum cDNA clone SCAGLR1021A08
            5', mRNA sequence.
ACCESSION   CA116904
VERSION     CA116904.1  GI:34970211
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
```

THIS PAGE BLANK (USPTO)